AUTONOMOUS MACHINE VISION FOR OFF-ROAD VEHICLES IN UNSTRUCTURED FIELDS

BY

QI WANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Biological Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Doctoral Committee:

Professor Qin Zhang, Chair
Associate Professor Eyal Amir
Associate Professor Tony E. Grift
Professor Alan C. Hansen
Associate Professor Lei Tian

ABSTRACT

A feasibility study of machine vision applications was conducted for agricultural vehicle navigation in open field environments, and focused on solving certain fundamental issues in vision-based agricultural vehicle navigation. Those issues were: (1) camera installation pose automatic calibration; (2) vehicle heading estimation and (3) field edge detection. A stereo color camera was selected to support the research on the three issues.

Stereo cameras have been used as perception sensors for agricultural vehicle navigation for years. One problem impeding their broader application is the difficulty of determining the camera's installation pose using conventional measuring tools, especially when they are used in an open-field agricultural environment. To solve this problem, an automated calibration method was developed to determine the camera's installation pose with respect to the vehicle frame. Based on this method, a binocular stereo camera acquired a sequence of images as the vehicle moved straight forward a short distance on relatively even ground. A machine vision algorithm was used to detect static feature points on the ground and track their three-dimensional (3D) motions in relationship to the vehicle. A plane fitting for the ground features was then used to determine the camera roll and pitch, and the tracked motions were used to estimate the camera yaw. The results obtained from the field test validated that this method was capable of determining the camera installation pose automatically in order to achieve a calibration accuracy of ±1 ° over approximately 10 m of test distance. The calibrated poses could be used to compensate for the navigation errors induced by the misalignment of the camera.

An image processing algorithm was developed to investigate the feasibility of using stereovision to estimate the heading direction of a moving vehicle in open agricultural field environments. The algorithm first detected, and then tracked, static natural ground features in every two consecutive images that were taken by a stereo camera mounted on a vehicle while the vehicle was in motion. These static features were used as references to calculate the vehicle's three-dimensional (3D)

motion. In the final stage, the vehicle heading direction was estimated using the 3D motion. Working with a series of sequential image frames taken while the vehicle was in motion, the algorithm continuously estimated the vehicle heading direction. Field tests were conducted to evaluate its usability. When the vehicle traveled straight forward, the proposed algorithm worked properly. When the vehicle traveled in an oscillating mode, the algorithm responded properly when the vehicle turned, but with less estimation accuracy than in the straight traveling mode. The field tests showed that it is possible to use stereovision to estimate a moving vehicle's heading direction in an open agricultural field.

Field edges are important references for human drivers who steer vehicles in agricultural operations. This research explored the possibility of using machine vision to detect field edges in open field agricultural environments. A detecting algorithm was proposed based on the hue difference between an open field and its grass-covered edge. Field tests showed that the algorithm was capable of distinguishing a relatively clear edge from an open field. However, when the field edge was not clear, the algorithm was unable to identify it due to the existence of noise. This research showed that images with lower resolution were less affected by noise. The same algorithm detected unclear field edges after reducing noise by lowering image resolution. Color change adaptability was also implemented in order to improve the algorithm's robustness. As a result, it was possible to use machine vision to detect the grass covered edges of an open agricultural field.

This research proved the feasibility of the machine vision applications in the three targeted problems, and has shown that machine vision is capable of navigating agricultural vehicles in open field environments.

ACKNOWLEDGEMENTS

I am heartily thankful to my supervisor, Qin Zhang, whose encouragement, guidance and support enabled me to truly understand the subject. His perpetual energy and enthusiasm for research motivated me onwards. In addition, he was always accessible and willing to help his students with their research. As a result, my research life became smoother and more rewarding. I would like to convey my special appreciation to other members of my committee, Eyal Amir, Tony E. Grift, Alan C. Hansen and Lei Tian for providing assistance and counsel during both my preliminary and final exams.

All of my colleagues at the Information Agriculture and Infotronic Systems Laboratory helped make it a congenial place to work. In particular, I would like to thank Francisco Rovira-Más, Michio Kise, Scott Dixon, Ryan Kingdon, Bhaskara Daddupaduvari and Yingjie Gao for their friendship and help during the past four years. Several other people also helped inspire me as a result of our interactions during the long hours in the lab. Thank you.

I would like to express my gratitude to my friends for their support, encouragement and sympathy. Weilin Wang, Muyuan Wang, Jonathon McCrady, Liangji Huang, David Oost, Haibei Jiang, Glen Menezes, Zhijun Yin, Haitao Xiang and Shu Zhang, it was great being the beneficiary of your friendship. Thank you.

My deepest gratitude goes to my parents for having been a permanent source of strength and affection. Their love, support, encouragement and advice are the most valuable things I possess in my life. Any achievements I may attain have been motivated by them, and this work is dedicated to them.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

During the past 100 years, mechanization has revolutionized agriculture, the most ancient human occupation. It has transformed the ways in which people are employed and produce their food. One form of evidence is that in 1900, farmers constituted 38% of the US labor force. By 2000, that number had plunged to 3% (Constable and Somerville, 2003). Since the 1990s, with the adoption of the concept known as "precision agriculture", some cutting edge technologies have changed the ways in which people operate agricultural machinery. For example, computer monitoring systems, global positioning systems (GPS) and self-steering programs have made tractors and implements more precise and more fuel-efficient. Currently, however, operating agricultural machinery remains repetitious labor, which leaves farmers exhausted after they spend long hours working in field operations. The same problem occurred during industrialization. Mechanization provided human operators with machinery to assist them with meeting the physical requirements of work, while still requiring extensive human intervention. Automation was a step beyond mechanization that greatly reduced the need for human sensory and mental requirements. This industrial process represented a possible solution that can be used in agricultural applications, even though agricultural environments have more unpredictable factors than industrial environments. Ting and Grift (2005) pointed out that automation is built on the success of agricultural mechanization, and is expected to have a huge impact on the future of agricultural and food systems and related environmental issues.

The purpose of automation is to equip engineering systems with perception, reasoning/learning, communication, and task planning/execution capabilities (Ting and Grift, 2005). During the 1980s, some innovators began to conduct research on the use of automation for farming purposes. Research achievements during the past two decades have shown the potential of automation for enhancing the capabilities of agricultural machinery. Machine vision has been used as a perception tool to guide autonomous tractors and combines in structured agricultural environments (Reid and Searcy, 1987; Billingsley and Schoenfisch, 1995; Pinto and Reid, 1998;

Benson et al., 2000a; Kise et al., 2005; Subramanian et al., 2006). During 1995, the first GPS guided autonomous tractor was created at Stanford University (O'Connor et al., 1995). Later on, multi-sensor guided tractors, which applied GPSs, cameras, sonar systems, geomagnetic direction sensors, and inertial measurement units (IMU), were introduced by researchers in Japan, the U.S. and Europe (Reid et al., 2000; Keicher and Seufert, 2000; Noguchi et al., 2002; Kise and Noguchi, 2005). Laser scanners, meaning a three-dimensional (3D) surrounding perception sensor, have been used for autonomous tractor navigation in structured agricultural environments as well (Barawid et al., 2007).

Along with those achievements, two challenging problems have emerged as subjects for further research in agricultural automation. First, most of the existing research has focused on structured agricultural environments, such as fields with crop lines and orchards with alleys. However, certain types of agricultural operations, such as plowing and seeding, take place in open fields. Therefore, autonomous navigation for tractors in open field environments should be explored. Leemans and Destain (2006) developed an algorithm for finding a guidance direction for a seeder designed to follow furrows using machine vision. That was the start for exploring this subject, but there remain a number of unsolved issues. Second, a number of published navigation methods for farming vehicles have relied on expensive sensors, such as laser scanners, high accuracy GPSs, and high accuracy IMUs. All of these methods have validated the idea of autonomous farming, but left room for discussion of the possibility of using relatively inexpensive sensors to achieve the same goal.

This dissertation is intended to explore the use of machine vision, an inexpensive means of perception, for tractor guidance in open agricultural fields. The overall goal was to investigate the feasibility of using machine vision to find guidance clues based on the natural visual references in an open agricultural field. To achieve this goal, a binocular stereo camera was selected, because it could provide both color information and 3D space information about a scene. This research was conducted to achieve the following three specific objectives:

Objective 1: developing an auto-calibration method for the purpose of determining camera installation pose for stereovision-based vehicle navigation.

Objective 2: investigating the feasibility of using machine vision for the purpose of determining the heading direction of a working agricultural vehicle in an open field without manmade references.

Objective 3: investigating the feasibility of using machine vision for the purpose of detecting field edges.

CHAPTER 2: LITERATURE REVIEW

This section has two parts: the first part is a review of autonomous navigation for agricultural vehicles; the second part is a brief review of recent research work regarding machine vision applications for ground vehicle navigation in off-road environments.

2.1 AUTONOMOUS NAVIGATION FOR AGRICULTURAL VEHICLES

2.1.1 Purpose of Navigation for Agricultural Vehicles

Agricultural machinery operation usually requires numerous repetitions. According to the US Census of Agriculture in 2002 (http://www.agcensus.usda.gov/), the average farm size in the United States was 441 acres. This number meant that a farmer would have to drive a tractor or combine back and forth many times in order to cover the entire field during a seeding or harvesting season. Intensive repetitions wear down people and cause mistakes, or even accidents, during operations. The automatic guidance of agricultural machinery can help relieve farmers from tedious repetitive field maneuvering and improve their productivity.

The rapid development of information technology has improved the probability of the introduction of the automatic guidance to agricultural vehicles. Several navigation methods, including the use of mechanical feelers, machine vision, GPS, geomagnetic direction sensors (GDS) and IMU, have been brought to agriculture (Reid et al., 2000; Keicher and Seufert, 2000).

2.1.2 Machine Vision Applications in Farming Vehicle Navigation

Machine vision applications for farming vehicle navigation can be divided into two categories: one category is applications in structured environments, such as fields with crop rows; the other one category is applications in ill-structured environments, such as open fields without crops. Since machine vision was introduced to agricultural automatic guidance in the mid-1980s, most

of the research has focused on applications in structured environments, while little research has focused on developing applications for ill-structured environments.

2.1.2.1 Application in Structured Environments

Crops have different periods of growth, and the crop appearances in different growth periods can be quite different. When crops simply sprout, they are typically spotty in the sort of fields that are shown in Figure 1. Therefore, a crop row is formed by non-continuous spotty plants. When crops grow taller, the flourishing canopy makes a crop row continuous as shown in Figure 2. During harvest seasons, ripe crops may not show up in the form of clear rows (Figure 3). Given the visual diversity present during a growing season, machine vision-based automatic guidance for agricultural vehicles can be classified into three categories:

(1) Fields with spotty crops.

(2) Fields with continuous crop rows.

(3) Fields with ripe crops.

Figure 1. A field with spotty crops.

Figure 2. A field with continuous crop rows.

Figure 3. A field with ripe crops.

Billingsley and Schoenfisch (1995) conducted research about how to find a guidance direction in a field filled with newly sprouted plants, such as cotton. They used brightness thresholding to differentiate crops from the background. Afterwards, they conducted linear regression to find crop rows. The movement of the vanishing point of these regression lines exhibiting a change in heading. Another approach was used to solve the same type of problem for different crops. Marchant (1996) used the Hough transform to detect crop rows of newly spouted cauliflower, sugar beets and wheat. This approach showed acceptable level of robustness for dealing with

problems such as missing plants and the existence of weeds.

As crops grow, the spotty plants in a row visually fuse to form a solid line. In order to develop machine vision-based automatic guidance for this field condition, Reid and Searcy (1987) found that crop canopy reflectance was significantly greater than background reflectance in the near-infrared (NIR) zone of the electromagnetic spectrum. They developed a dynamic thresholding algorithm based on the distribution of NIR image pixels to segment crop rows found in a natural scene. In order to find a guidance direction in a row segmented image, Pinto and Reid (1998) explored the application of the principle component analysis method. Later research projects dealt with images in the visible spectrum and successfully determined a guidance direction. For example, the vision-based guidance system developed by Han et al. (2004) included row segmentation by using a K-means clustering algorithm, row detection by using a moment algorithm, and guidance line selection by using a cost function. Rovira-Más et al. (2005) applied the Hough transform and connectivity analysis to process images of a vehicle's forward view and then used them to determine an appropriate pathway in a field. Stereovision was also an approach used to detect crop rows. Kise et al. (2005) used stereo images to analyze the geometry of field terrain profiles, and to detect crop rows and navigation points.

Many fields do not usually show clear crop rows during harvest operations. As regards machine vision-based automatic guidance, the features of interest are no longer multiple crop rows, but rather the edges between the cut and uncut crops. Ollis and Stentz (1997) used color segmentation to detect the cut and uncut edges in hay fields. Their method successfully guided a harvester over 60 acres at a speed of 7 $km \cdot h^{-1}$. As regards other crops, such as corn, the algorithm developed by Benson et al. (2000b) used two-class K-means thresholding to segment and classify the points as cut or uncut crops, which showed acceptable robustness when guiding a corn harvester.

2.1.2.2 Application in Ill-Structured Environments

In an ill-structured environment such as an open field (Figure 4), there are not as many significant visual clues as is the case in a structured environment to provide the necessary references for finding a guidance direction. Few researchers have used machine vision for such applications. Leemans and Destain (2006) used the Hough transform to detect furrows in an image, which was then used to determine a guidance direction for a seeding vehicle. That was the beginning of research on machine vision-based automatic guidance in ill-structured environments. Meanwhile, given that agricultural field environments are subject to changes in natural conditions such as illumination and moisture, the reliability of the present seeding direction determination method deserves more exploration. Certain other types of visual clues, such as color, texture and elevation, could be considered as possible means of augmenting the reliability of their method.

Figure 4. Seeding operation in an open field.

2.1.2.3 Summary

The literature review revealed that the existing research on machine vision-based automatic guidance has primarily focused on applications in structured environments. As regards the applications in ill-structured environments, very little research has been reported. Therefore, there remain many challenges that will need to be addressed concerning agricultural machinery

automatic guidance in ill-structured fields.

2.2 MACHINE VISION APPLICATION TO GROUND VEHICLE NAVIGATION IN OFF-ROAD ENVIRONMENTS

Machine vision has been an active area within robotics research for several decades. Its application to ground vehicle navigation in off-road environments can be traced back to the 1970s, when researchers (Gennery, 1980; Moravec, 1983) working at Stanford University applied stereovision to mobile robot navigation. The Stanford Cart (Figure 5) was a mobile robot platform, according to the reported research. It used stereovision to locate surrounding objects in 3D space and deduced its own motions. Due to the computation limitations of hardware at that time, the system was reliable for short runs, but operated quite slowly (Moravec, 1983).

During the 1980s, Carnegie Mellon University (CMU) and the NASA Jet Propulsion Laboratory (JPL) were in the forefront of research on mobile robots. Some of their achievements in the applications of machine vision in their work laid the foundation for the rapid progress of vision-based ground vehicle navigation in later years. After the work on the Stanford Cart, researchers at CMU developed a new mobile robot platform called CMU Rover that addressed some of the Stanford Cart's problems, such as the weak computational capability (Moravec, 1983). The improvements included hardware upgrades, as well as new algorithms for perception using stereovision. One highlight of the work was a stereovision-based visual odometry algorithm that produced the first quantitatively accurate vision-based egomotion estimation results (Matthies and Shafer, 1987) and led to visual odometry algorithms that are used for certain ground vehicle navigation projects in certain types of off-road environments on Earth (Nister et al., 2006; Agrawal et al., 2007) and in NASA's Mars Exploration Rover (MER) project (Cheng et al., 2006).

Figure 5. The Stanford Cart (http://www.stanford.edu/~learnest/cart_files/image008.jpg).

A later mobile robot platform called Navlab (Figure 6), which was developed at CMU, was a robot vehicle that has the computational power to perform onboard terrain perception, terrain mapping and path planning (Matthies et al., 2007). It used a monocular color camera for road-following on structured roads. For tasks conducted in ill-structured environments, such as cross-country navigation, the robot needed both information derived from appearance (such as road location in a color image, or terrain type), and the geometry of the observed environment. The Navlab team selected an active sensor: a laser range scanner rather than a passive solution, such as stereovision, for collecting 3D terrain appearance data, due to the drawbacks of a passive solution at that time, which included high computational demand, difficulty in ranging bland surfaces, and reliance on ambient lighting (Thorpe et al., 1991). However, during the late 1980s, Matthies' work at JPL made a research breakthrough with real-time, area-based stereovision algorithms that facilitated the first stereovision-guided autonomous, off-road traverse (Matthies, 1992). Since then, stereovision has become a competitive 3D perception approach in the area of autonomous ground vehicle navigation.

Figure 6. The first Navlab mobile vehicle developed at CMU (http://www.cs.cmu.edu/afs/cs/Web/Groups/ahs/images/navlab_1_5_images/navlab1.color.small.gif).

The 1990s was a period during which mobile robot vehicles began to approach real-time autonomous navigation. The onboard hardware began to feature more powerful computing that made it possible to complete more complex algorithms in a short period of time. CMU's machine vision-based robot vehicle, Navlab 5, traveled 2,849 miles from Pittsburgh to San Diego with 98.2% of its driving conducted autonomously (Pomerleau and Jochem, 1996). As regards the stereovision applications, Konolige's stereo algorithm could run at up to 30 Hz for 320×240 imagery using DSPs (Konolige, 1998). Given such computational power, hardware is no longer a bottleneck for stereovision.

During the 2000s, several off-road autonomous vehicle projects brought machine vision applications for navigation to a higher level. The Demo III program (Figure 7) was funded by the Army Research Lab (ARL) and the DARPA Perception for Off-Road Robotics (PerceptOR) program, and investigated off-road navigation in tough terrain environments, which included terrains such as ditches, grass, water, rocks and trees (Bellutta et al., 2000; Krotkov et al., 2007).

Figure 7. Demo III autonomous off-road vehicle (http://www.gdrs.com/news/gallery/image.asp?image_id=6).

Stereovision played an important role in NASA's MER project. Figure 8 shows one of the rovers. Rover navigation was conducted using three sets of paired stereo cameras: one pair of hazard cameras looked forward under the solar panel in the front, while a second pair of hazard cameras looked backwards under the solar panel in the rear, and a third pair of navigation cameras was mounted on the pan-tilt mast (Matthies et al., 2007). Stereovision achieved two basic functions of navigation: visual odometry and path planning. There is no absolute positioning system such as GPS on Mars, so stereovision-based visual odometry is a relative positioning method that can help rovers locate themselves (Cheng et al., 2005). Stereovision-based path planning software can percept terrain (Biesiadecki and Maimone, 2006) and predict slippage on slopes (Angelova et al., 2007).

Figure 8. NASA's MER rover (http://www.jpl.nasa.gov/images/mer/2003-02-26/PIA04413-med.jpg).

Similar types of research projects for off-road mobile robot navigation have been tested for use on Earth as well. Researchers have used visual odometry based on either monocular vision or binocular vision, as a positioning method. Nister et al. (2006) worked on a ground vehicle system that was capable of estimating motions using input from both monocular and stereo videos. The most reliable, accurate, and conveniently applicable results occurred when they used the system in the stereo setting. Konolige et al. (2008) focused on computational efficiency for stereovision-based motion estimation, and developed a real-time system for tracking the motions of a mobile robot moving at about 1 $m \cdot s^{-1}$.

Agricultural open fields are a kind of off-road environment with their own special conditions. These fields typically have boundaries (field edges), and the ground surfaces within these boundaries are relatively even when there are not a number of obstacles (Figure 9). However, there have been few reports regarding vehicle motion estimation using machine vision in the agricultural area.

Figure 9. The top view of some fields on the South Farms of the University of Illinois (Urbana, IL, USA) (http://maps.live.com).

CHAPTER 3: AUTO-CALIBRATION METHOD TO DETERMINE CAMERA POSE FOR STEREOVISION-BASED NAVIGATION

3.1 INTRODUCTION

Stereo cameras have been used as perception sensors of surroundings for agricultural vehicle navigation during the last ten years. Their capabilities for depth perception provide more flexibility in 3D sensing than is the case for other 3D perception sensors. Although stereovision uses a great deal of computational power, significant advances in electronics and processor speed have led to the enhancement of its benefits and the amelioration of its disadvantages (Rovira-Más et al., 2009). An interesting application of stereovision for ground vehicle navigation purposes is the determination of the heading direction in agricultural field operations, which seeks to steer the vehicle to follow a desired path by tracking the positions of detected landmarks that are relevant to the moving vehicle. Therefore, the camera pose (roll, pitch and yaw) with respect to the vehicle becomes critical information for obtaining navigation commands from visual features. In stereovision-based agricultural vehicle navigation, a common configuration of the camera is to install the camera have it look forward with a pitch angle pointing downwards at the ground. In this case, however, it is not easy to determine the pose using conventional measuring tools such as a spirit level and a protractor, particularly in the open fields that often comprise agricultural environments. Furthermore, the camera pose may change during field operations. Because the calibration process is too complicated for the average user and is beyond the ability of most drivers to perform in the field, an automated calibration process would be quite helpful if it could be made available prior to the release of a practical stereovision-based navigation system to farmers.

Similar problems have occurred with stereovision-based vehicle navigation in on-road environments. An important difference when comparing on-road applications with off-road applications is that on-road environments are well-structured, and have edges that are normally well indicated by lane markers and road curbs. Such boundaries are often selected as calibration

references to estimate the installation pose of a stereo camera (Franke et al., 1998; Coulombeau and Laurgeau, 2003; Labayrade et al., 2003; Collado et al., 2006; Marita et al., 2006; Lamprecht et al., 2007). Plane fitting for road surface estimation is also used to support camera positioning (Weber and Atkin, 1997; Sappa et al., 2006). When the scenario is changed from on-road environments to off-road or agricultural environments, there is little research available regarding this issue. The major challenge for off-road applications is that ill-structured terrain introduces more uncertainties into the calibration process. For example, in the planting operation in Figure 10, the field in front of the tractor does not have significant landmarks or path indicators to follow, and the texture of crop residue often changes without providing regular patterns that can be followed.

This research copes with these challenges that are found in open field agricultural environments and proposes an automated calibration method for estimating the installation pose of a stereo camera for vehicle navigation.

Figure 10. The field in front of a planting tractor does not have significant patterns.

3.2 THE CONCEPT OF CAMERA POSE AUTOMATIC CALIBRATION

3.2.1 System Architecture

The experimental vehicle platform used for this research was the John Deere Gator™ Utility Vehicle shown in Figure 11 (Deere & Co, Moline, IL, USA). A binocular stereo camera manufactured by Videre Design (Menlo Park, California, USA) was mounted on a rigid frame in the front part of the vehicle. An onboard computer was used to support the camera pose auto-calibration software.

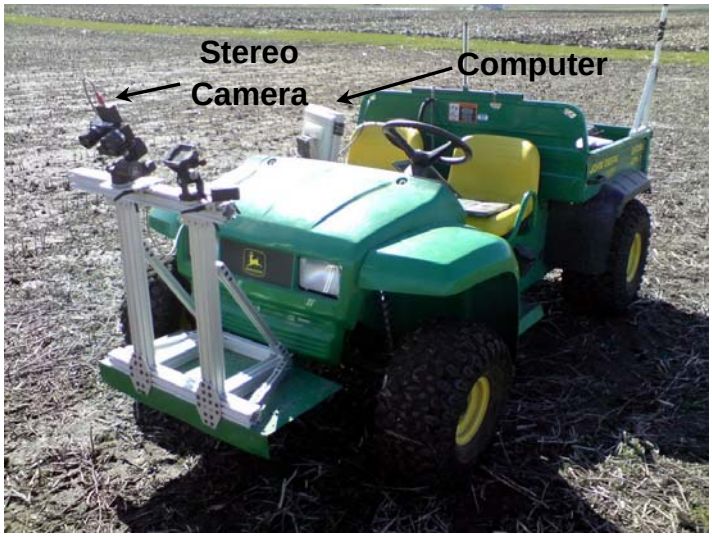


Figure 11. Experimental vehicle: John Deere Gator™ Utility Vehicle.

3.2.2 Definition of Coordinate Systems

Before discussing the stereo camera installation pose on agricultural vehicles, it is crucial to define the coordinate systems employed in this research. There were three coordinate systems: the camera coordinate system $X_cY_cZ_c$, the vehicle coordinate system $X_vY_vZ_v$ and the ground coordinate system $X_gY_gZ_g$ (Figure 12). They are defined as follows:

The origin of the camera coordinate system was located at the optical point of the left lens of the binocular camera (Figure 13), with the Z axis (Z_c) pointing forward and perpendicular to the left imager (camera sensor), the X axis (X_c) pointing to the right and the Y axis (Y_c) pointing

downwards, and both were parallel to the image plane.

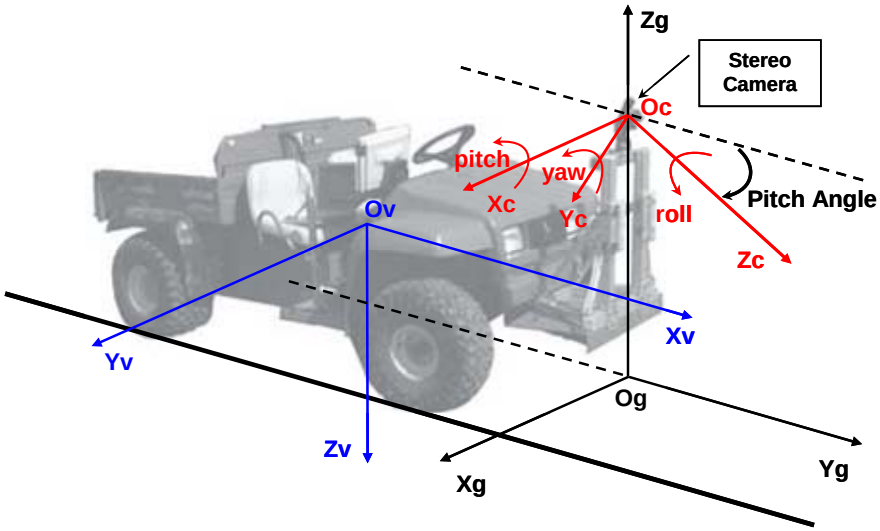


Figure 12. Definition of coordinate systems.

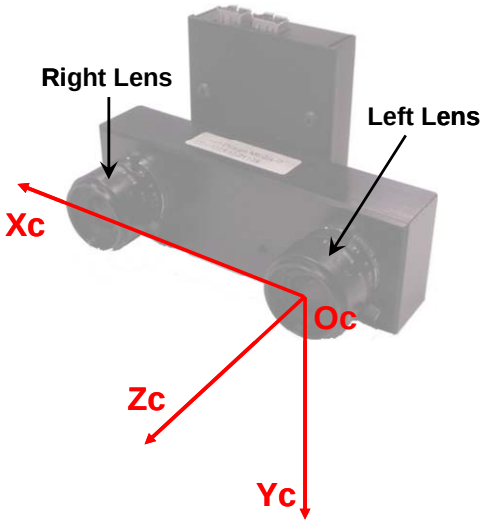


Figure 13. Camera coordinate system.

The origin of the vehicle coordinate system was at the center of gravity of the vehicle platform, with the X axis (X_v) pointing to the front, the Y axis (Y_v) pointing to the right and the Z axis (Z_v) pointing downwards, as represented in Figure 12.

The ground coordinate system was defined at the starting point of the navigational path, and did not change when the vehicle traveled. The origin O_g was at the intersection of the ground surface

and the plumb line that went through the origin O_c. The Y axis (Y_g) pointed to the front of the vehicle, the X axis (X_g) pointed to the right and the Z axis (Z_g) pointed upwards, as illustrated in Figure 12.

3.2.3 Camera Installation Pose

When a stereo camera is mounted on a vehicle, its orientation with respect to the vehicle, which is called *pose* or *attitude*, is critical for stereovision-based navigation. The reason for this is that the visual clues detected by the stereo camera lie in the camera coordinate system, but they must be mapped into the vehicle coordinate system or the ground coordinate system before they can be used for vehicle navigation. The pose can be completely described using the roll, pitch and yaw angles of the camera (Figure 12).

As mentioned above, a common installation pose for cameras that are implemented in agricultural vehicles sets the camera looking forward with a pitch angle pointing downwards at the ground (Figure 12). Ideally, the camera should have only the desired downward pitch angle (Figure 14). However, in a real installation, it is inevitable for the camera to possess some roll and yaw angles (Figure 15) because the conventional measuring tools for estimating the camera orientation are insufficiently accurate to ensure a perfect pose in the vehicle coordinate system. Therefore, the mounted cameras must unavoidably suffer some pose errors, including some small roll, pitch and yaw angle errors with respect to the vehicle frame. Such pose errors reduce the navigational accuracy of agricultural vehicles that are based on a 3D visual sensor.

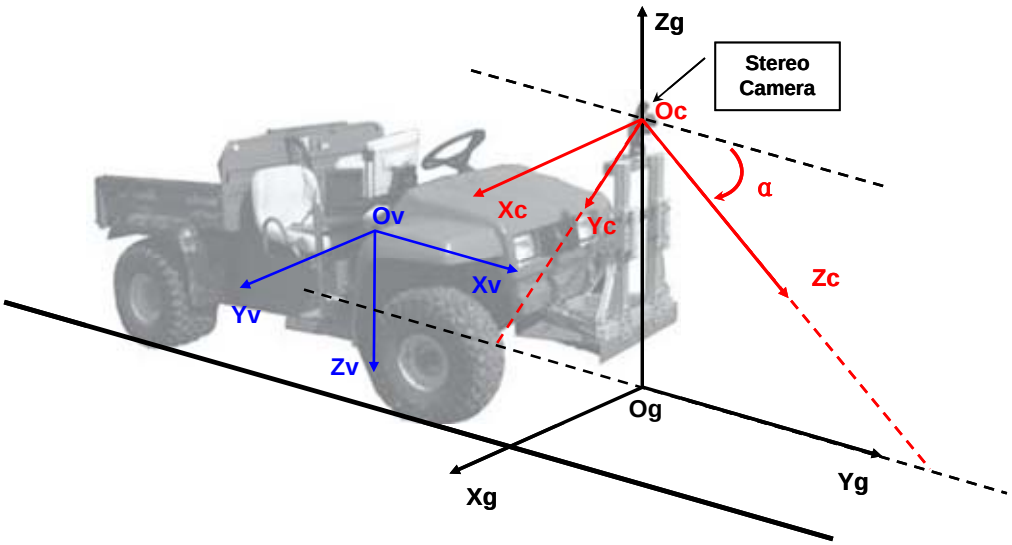


Figure 14. A camera installed with only a pitch angle α.

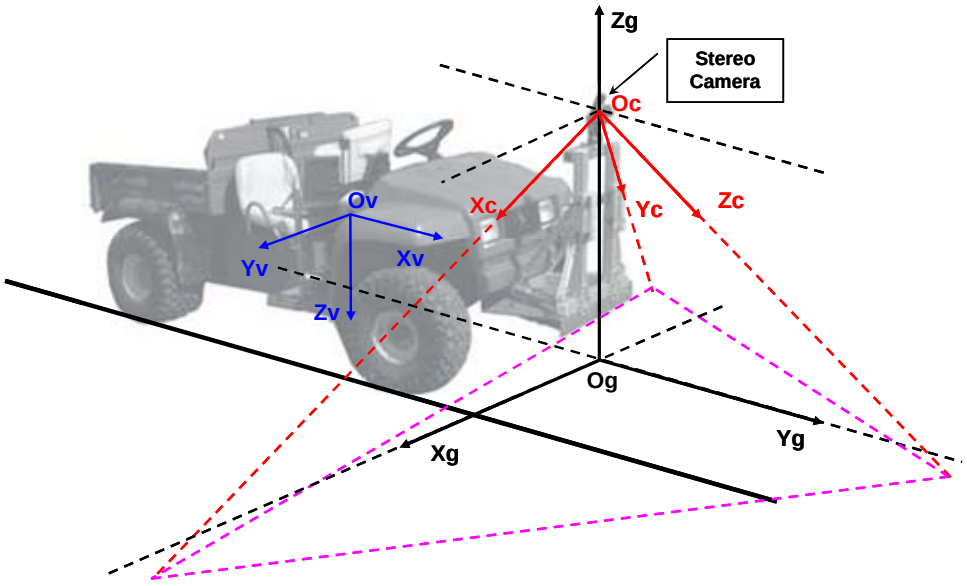


Figure 15. Camera installation with roll, pitch and yaw.

3.2.4 Installation Pose Calibration

The determination of the camera's roll, pitch and yaw with respect to the vehicle is important in navigation applications. However, the geometric relationship between the camera coordinate system and the vehicle coordinate system is often difficult to determine with fully acceptable

accuracy because the origin O_v (the vehicle's center of gravity) is often difficult to locate. Note that when a vehicle is on a flat ground surface, meaning when the axes satisfy the relationships: X_v is parallel to Y_g, Z_v is parallel to Z_g and Y_v is parallel to X_g as shown in Figure 12, it is possible to use the ground coordinate system instead of the vehicle coordinate system for the purpose of determining the roll, pitch and yaw angles of the camera. The assumption being made for this replacement is that the vehicle is on even terrain and that the camera is facing towards a point on the ground that is some distance away in the front of the vehicle. According to the definition of the three coordinate systems used, when the camera has no roll, pitch or yaw, there are three equivalence relations:

(1) If roll = 0, the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gY_g$ (ground surface).

(2) If pitch = 0, the coordinate plane $X_cO_cZ_c$ is parallel to the coordinate plane $X_gO_gY_g$.

(3) If yaw = 0, the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gZ_g$.

These three equivalence relations lead to a three-step automatic pose calibration. The basic idea of this pose calibration method is to rotate the camera coordinate system from an arbitrary installation pose to an ideal pose without roll, pitch or yaw (Figure 16). The rotating angles obtained along the axes Z_c, X_c and Y_c indicate the camera roll, pitch and yaw, respectively.

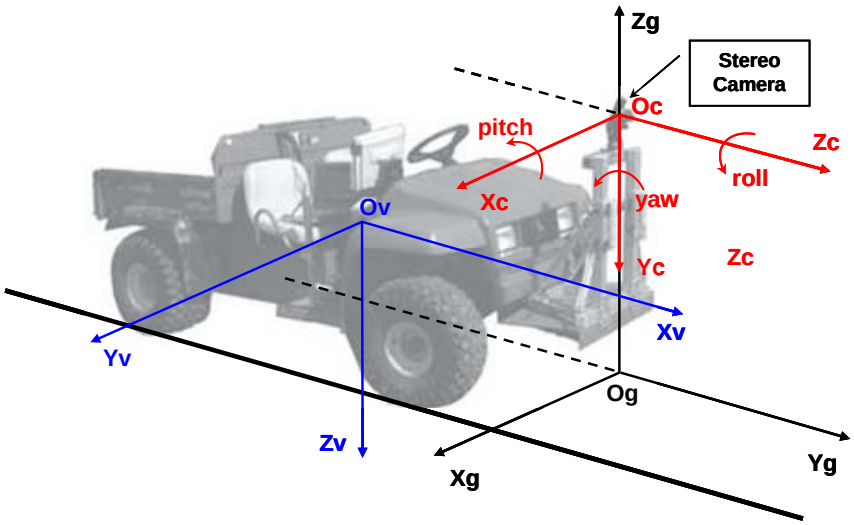


Figure 16. Camera pose without roll, pitch or yaw.

The pose determination method proposed follows the order of roll, pitch and yaw:

(1) As shown in Figure 17, a rotation along the axis Z_c is used first in this procedure in order to find the camera roll. The camera coordinate system is rotated around the axis Z_c until the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gY_g$. The opposite number of the rotation angle is the roll angle of the camera, β.

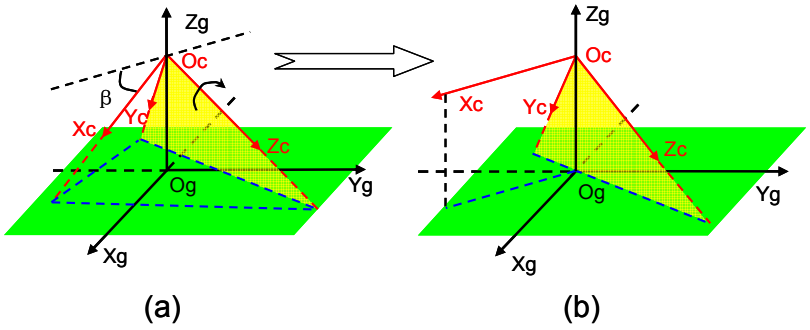


Figure 17. A rotation around the axis Z_c for the purpose of finding the camera roll.

(2) A rotation around the axis X_c follows in order to find the camera's pitch (Figure 18). The camera coordinate system is rotated around the axis X_c until the coordinate plane $X_cO_cZ_c$ is parallel to the coordinate plane $X_gO_gY_g$. The opposite number of the rotation angle is the pitch angle α.

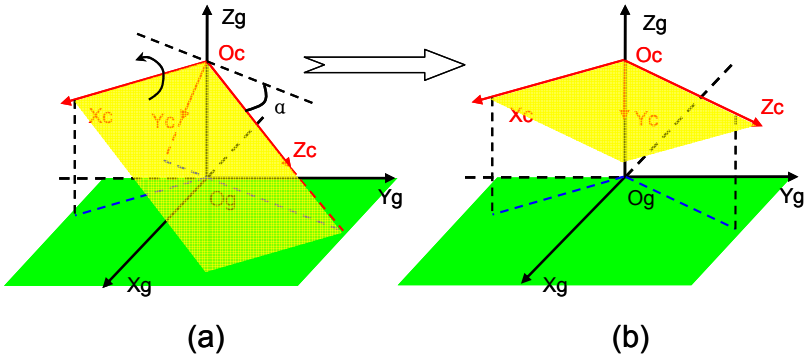


Figure 18. A rotation around the axis X_c for the purpose of finding the camera pitch.

(3) Finally, a rotation around the axis Y_c determines the camera's yaw (Figure 19). The camera coordinate system is rotated around the axis Y_c until the coordinate plane $Y_cO_cZ_c$ is perpendicular

to the coordinate plane $X_gO_gZ_g$. The opposite number of the rotation angle is the yaw angle φ.

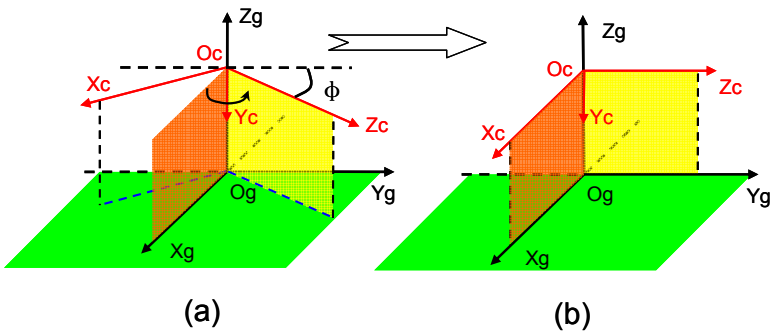


Figure 19. A rotation around the axis Y_c for the purpose of finding the camera yaw.

3.3 ALGORITHM DESIGN

3.3.1 Overview

This section introduces the design of a machine vision algorithm that was used to conduct the proposed three-step camera pose auto-calibration method.

In calibration steps 1 and 2, the reference plane used to determine roll and pitch angles is the coordinate plane $X_gO_gY_g$, which is actually the ground surface. Therefore, if the ground surface can be represented in the camera coordinate system, then both the roll and the pitch angles can be calculated. In calibration step 3, the reference used to determine the yaw angle is the coordinate plane $X_gO_gZ_g$, which may not be a user-friendly reference in machine vision applications because there is no constant feature that can be visualized. However, it is useful to define the camera yaw angle, which is φ in Figure 20. This figure describes the coordinate systems after roll and pitch corrections have been carried out, and it is equivalent to transforming the camera coordinate system to a pose with yaw only. In such a case, when the vehicle moves straight forward on a flat ground, its motion direction as estimated in the camera coordinate system is an indicator of the camera yaw. Figure 20 also shows that the angle between the vehicle motion direction and the axis

Z_c is φ_v, which equals φ. Therefore, if the vehicle motion direction can be measured in the camera coordinate system, the camera yaw can be determined.

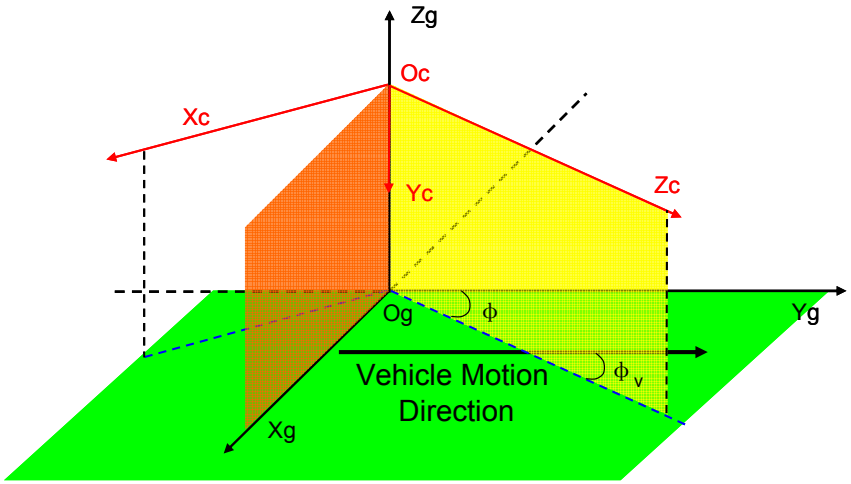


Figure 20. Using the vehicle motion direction to estimate the camera yaw angle.

The method for camera pose auto-calibration can be summarized in terms of the following two tasks: (1) determining a flat ground surface in the camera coordinate system for roll and pitch calibration; and (2) finding the vehicle motion direction in the camera coordinate system for yaw calibration.

In order to comply with the flat ground assumption, this calibration needs a piece of relatively even ground on which the vehicle can travel. As regards the first task (ground surface determination), it is necessary to know the location of at least three ground points in the camera coordinate system and then use plane fitting techniques to find the equation for ground surface. As regards the second task (determination of vehicle motion direction), several static visual landmarks were used as references in order to estimate vehicle motion. Unlike on-road environments, open agricultural fields are ill-structured; therefore, it is difficult to find distinct landmarks such as traffic lines or road curbs. However, due to the existence of crop residue, open fields are rich with visual textures, and such textures can be used as landmarks because they are detectable and trackable by machine vision algorithms. On the other hand, visual textures lie on the ground, so they are helpful for the

purpose of forming the ground points that can be used to estimate the ground surface. Therefore, ground textures were used as static features in order to deal with both of the tasks.

Based on selected ground features, a machine vision algorithm (Figure 21) was developed to fulfill the two tasks (ground surface determination and estimation of motion direction). During the calibration tests, the stereo camera shot images continuously while the vehicle moved straight forward on a relatively even ground for approximately 2 m. The image processing algorithm then identified and tracked static ground features in several pairs of consecutive image frames. The ground features which had been identified and tracked were first used to estimate the ground surface, which then led to the estimation and correction of the roll and pitch. Second, these tracked features and their relative motions between two consecutive image frames were used to estimate the vehicle motion direction, and then the camera yaw angle was calculated.

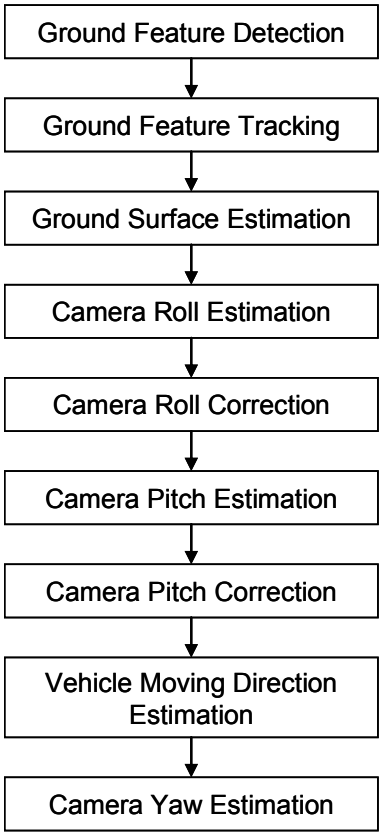


Figure 21. Process for stereo camera pose calibration.

3.3.2 Ground Feature Detection

As mentioned above, the texture of the ground was used to find static features. The left image in Figure 22 shows the texture of an open field with soybean residue. The difference in the colors between the soil and the randomly distributed residue generates ground textures. When observed from a distant perspective, this ground texture appears homogeneous, so it is difficult to determine a rule for converting its pattern into useful ground features. However, when looking at local detail, features can be readily recognized on the basis of textures. The right image of Figure 22 shows that the small areas pointed at by the three arrows do not look exactly the same, but rather show distinct corner shapes. It is fairly common to find this sort of corner features in textures produced by brightness contrast, as happens in agricultural fields with residue. Therefore, these corner features were used as static ground features in this dissertation.

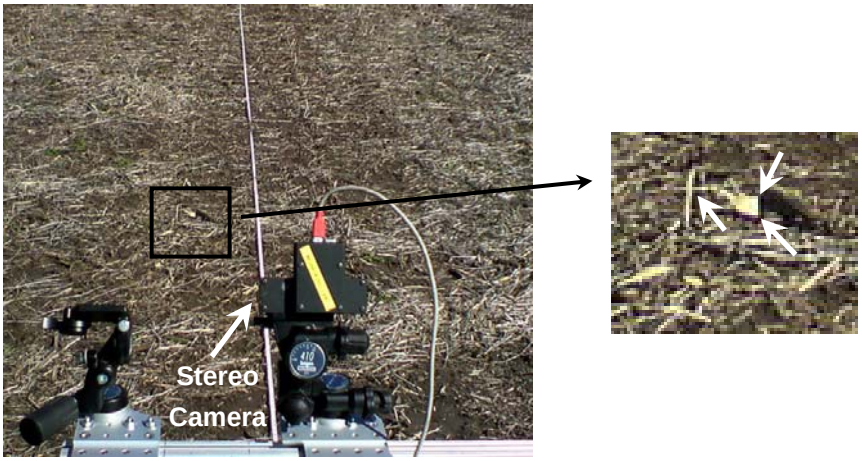


Figure 22. Texture in open fields with soybean residue. Left image: testing vehicle with stereo camera in open fields with soybean residue. The white straight line crossing the middle is an artificial reference used by the driver to steer. Right image: zoom-in view of the small piece of ground in the rectangle of the left image. Visual corners are pointed out by three arrows.

The Harris corner detector (Harris and Stephens, 1988) was used to identify corner features. Figure 23 shows that corner features are detected in a window at the lower part of the image. The lower part of the image is closer to the stereo camera than the upper part, so more accurate distances can be estimated using the camera, which is important for 3D reconstruction.

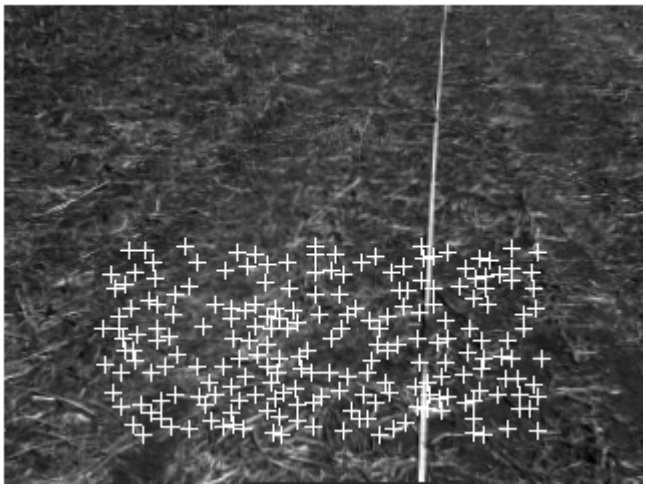


Figure 23. Corner feature detection results. Each cross marks a corner feature.

3.3.3 Ground Feature Tracking

3.3.3.1 Feature Matching

After corner features have been detected, the next step is to track them across two consecutive image frames. Corner features are marked as points (pixels), which are difficult to match across images. Because an image has a large number of pixels (76800 pixels for a 320×240 image), and each pixel can take on a value between 0 and 255 (for gray scale images), there are far too many pixels in one image to match an individual pixel in the following image. One way to solve this problem is to measure both the brightness of an individual pixel, and the brightness of a small window around the point of interest (Ma et al., 2005). Once these windows have been defined, corner (point feature) matching becomes window matching. Correspondence among windows across two images is calculated in order to determine the best matching window from one image to a window in the next image.

Because features move within the field of view when the camera moves, and the image plane is not parallel to the ground surface, there will be deformation of features and their surrounding areas between two consecutive image frames, which can lead to window mismatching. However, because the vehicle's speed is limited (between 0.5 and 1 $m \cdot s^{-1}$) during calibration, and also because the camera captures images at a speed of four frames per second, the deformations

between two consecutive images are small. For the purpose of simplifying the algorithm, it was assumed in this study that there was no feature deformation. In this instance, the sizes of matching windows for two consecutive images are the same.

In this algorithm, the feature points found in two consecutive image frames are detected separately. The attempt is made to match each feature in one image with every feature within a fixed mapping distance of it in the next image (Nister et al., 2006). The longer the mapping distance is, the more computational power is needed in order to perform the match. Therefore, selecting the proper distance is important for successful matching while maintaining computational efficiency. When the camera moves rapidly, the mapping distance becomes longer, and vice versa. In this research, the mapping distance was set at 50 pixels (for images with a resolution of 320×240 pixels), a distance that was sufficiently long enough to cover the majority of image disparities in this application.

The Normalized Cross-Correlation (NCC) method was used to determine the degree of similarity of the matching windows across two images during the matching process (Ma et al., 2005). The size of the matching window is typically between 5×5 and 21×21 pixels in the NCC method. Larger windows result in increased robustness at the price of increased computational cost. The current algorithm used an 11×11window as a tradeoff between robustness and efficiency. The windows in two images that produce the highest normalized correlation are the preferred matching windows.

An example of feature tracking obtained using the NCC method is shown in Figure 24. Most of the feature pairs show homogenous orientations, which indicate the motion direction of the camera. However, there are several matches shown in the image as well that are obviously incorrect. Filtering measures will be used to discard these outliers in the subsequent steps of the algorithm.

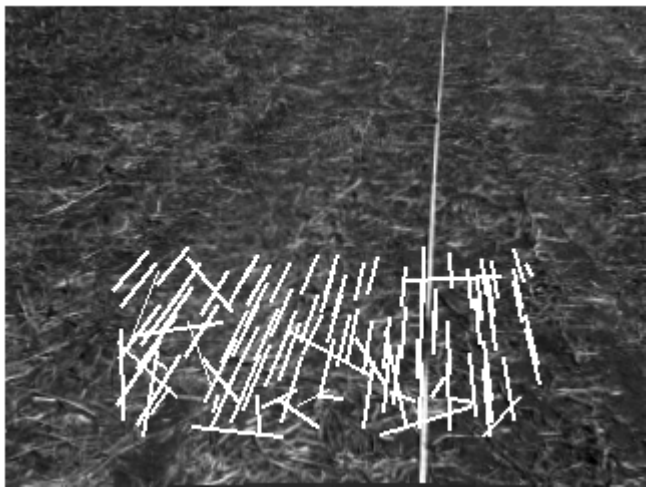


Figure 24. Feature matching results. All of the matched feature pairs between two consecutive image frames are connected by lines. The background is the first image frame.

3.3.3.2 Removal of Outliers

The Random Sample Consensus (RANSAC; Fischler and Bolles, 1980) has proven effective for the purpose of eliminating outliers from feature matching results (Cheng et al., 2005; Nister et al., 2006; Konolige et al., 2008). This calibration method used the RANSAC algorithm to refine the feature matching results in two stages:

Stage 1:

According to epipolar geometry, for all of the feature points, there is one fundamental matrix that can be used to transform their two-dimensional (2D) positions in the first image to the second image. Outliers were detected by fitting a fundamental matrix to the matched feature pairs using RANSAC.

Stage 2:

Given that the camera is a stereo camera, the results after removing outliers in step 1 can easily be transformed from the 2D image space to the 3D camera coordinate system. In the camera coordinate system, all of the feature points should travel the same distance between two consecutive image frames when the vehicle moves straight forward on a relatively flat surface (without significant rotations). Therefore, outliers were detected again by fitting a traveling distance to the matched feature pairs using RANSAC.

In most cases, outliers can be removed after the application of the two-stage RANSAC process. In order to ensure a successful calibration, this process was conducted twice in the algorithm. Although repetitions require additional computational power, given that the current calibration method is an offline approach, it is acceptable to use additional time in order to obtain better feature tracking results. Figure 25 shows that the outliers have been removed successfully from the image in Figure 24 after using the two-stage RANSAC process twice.

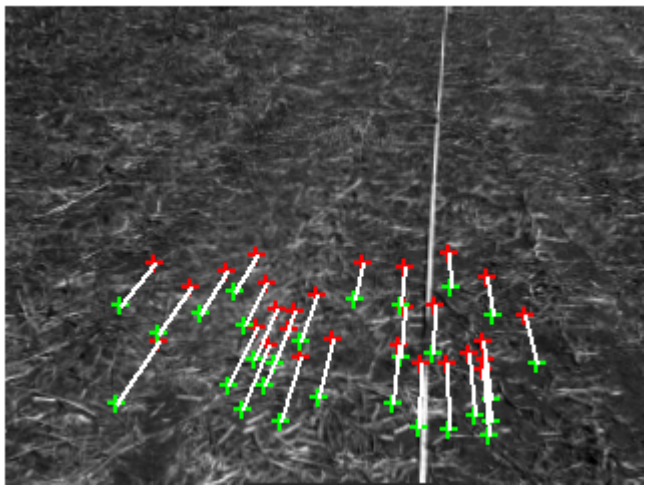


Figure 25. Feature tracking results after removing the outliers using the two-stage RANSAC process.

3.3.4 Ground Surface Estimation

The goal of specifying the ground surface is to determine the surface equation in the camera coordinate system. By taking advantage of stereovision capabilities, the ground surface can be readily transformed from a 2D image space to a 3D camera space. Figure 26(a) depicts the 3D point cloud of the relatively flat ground surface shown in Figure 25. When observed from a side view (Figure 26 (b)), the point cloud forms a flat surface in the camera coordinate system, and this flat surface can be represented by fitting a plane. However, not all of the points in the cloud are needed in order to fit the plane because:

(1) A plane can be determined by having as few as three points lying on it; too many points require additional computational power.

(2) The stereo camera accuracy degrades as perception range increases. This problem can be

observed in Figure 26 (b), as the point cloud varies its inclination at the far end from the camera, even though the actual ground was relatively flat.

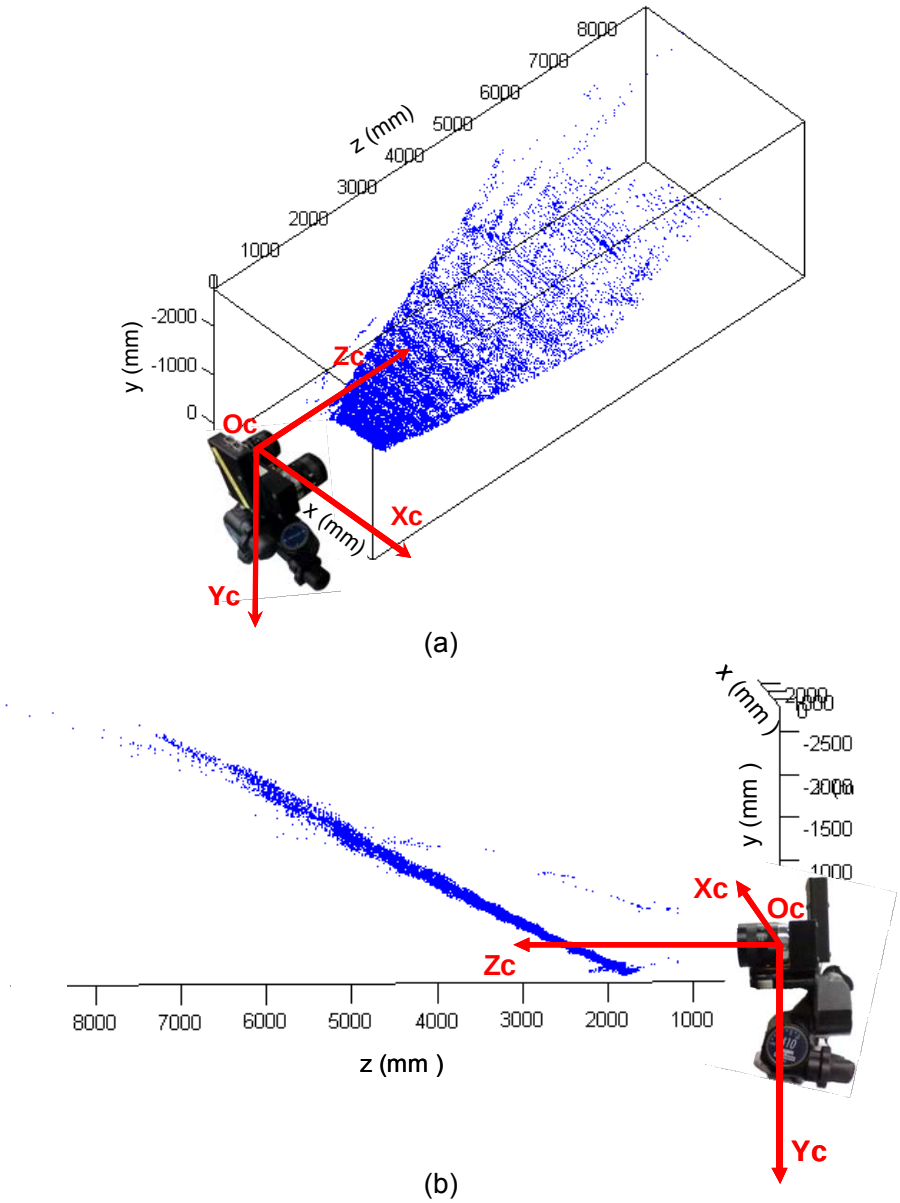


Figure 26. 3D point clouds of the ground surface in the camera coordinate system.

Therefore, those tracked ground features, which were lying at a close range from the camera and were limited in quantity, were selected for plane fitting. The MATLAB (2007a, The MathWorks, Natick, MA, USA) statistics toolbox was used for plane fitting by applying an orthogonal

regression. Figure 27 shows the plane fitting result, in which the ground surface was estimated by using the plane of Equation 1.

$$0.1184(x - 209.6) - 0.9042(y-43) - 0.4104(z-2458.4) = 0 \qquad [1]$$

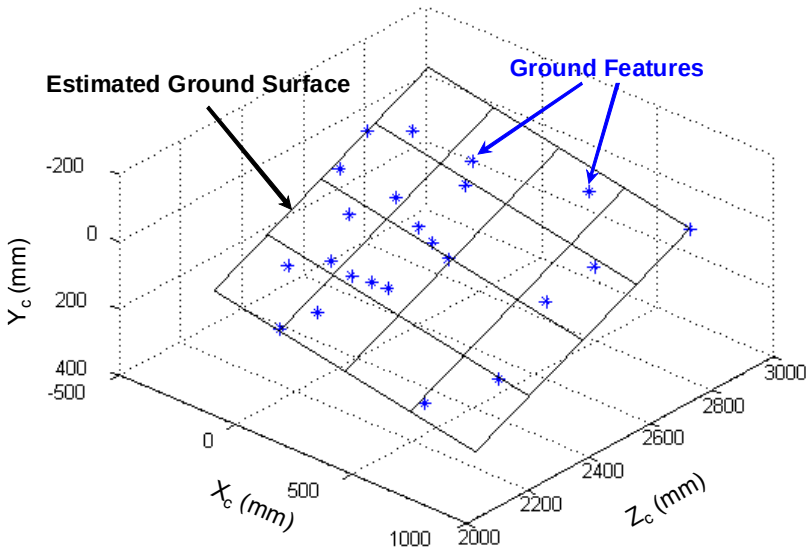


Figure 27. Ground surface estimation with plane fitting.

3.3.5 Camera Roll Estimation

As mentioned earlier, if the camera coordinate system rotates around the axis Z_c until the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gY_g$ (ground surface), the opposite number of the rotation angle is the roll angle β (Figure 17).

The ground surface estimation provides the equation for the coordinate plane $X_gO_gY_g$ in the camera coordinate system, and the angle between the coordinate plane $Y_cO_cZ_c$ and the coordinate plane $X_gO_gY_g$, which is named γ, can be calculated (Figure 28). According to the definition of the dihedral angle, γ is actually the angle between the plane $X_gO_gY_g$'s norm vector (axis Z_g) and the plane $Y_cO_cZ_c$'s norm vector (axis X_c). Figure 28 shows that the process of roll correction is performed by rotating the camera coordinate system around the axis Z_c until γ reaches 90°. Because γ can be either an obtuse angle (Figure 28(a)) or a sharp angle (Figure 28(b)), the

rotational directions for these two cases would be different for an acute approach to perpendicularity. When γ is an obtuse angle, the rotation direction is negative; when γ is a sharp angle, the rotation direction is positive.

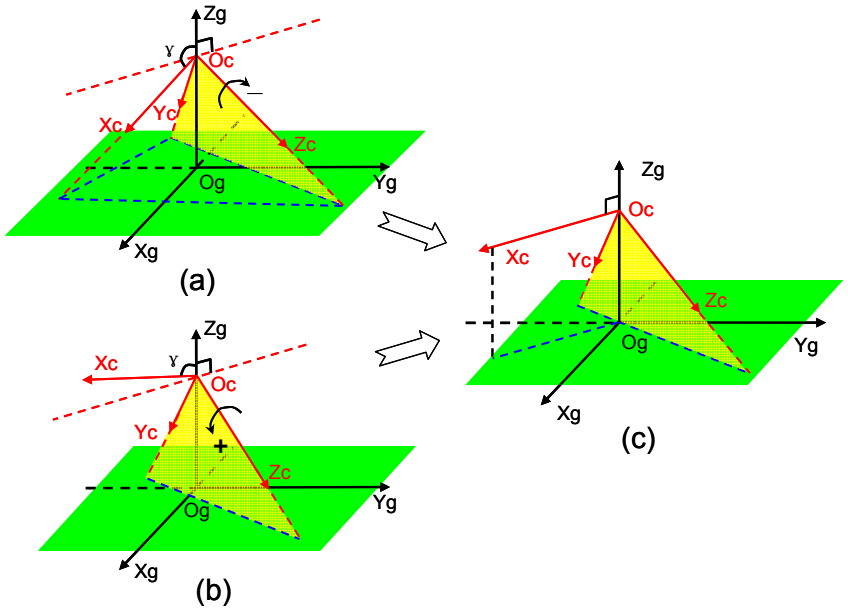


Figure 28. Camera roll calibration by rotating the camera coordinate system O_c along the axis Z_c. (a) Camera with a positive roll angle; (b) camera with a negative roll angle; (c) camera without roll angle.

The angle θ, defined as θ = 90° - γ, is used to determine how far the two coordinate systems are from being perpendicular. When the coordinate plane $Y_cO_cZ_c$ is perpendicular to the coordinate plane $X_gO_gY_g$, θ equals 0. Therefore, θ is constantly checked as a control parameter during a rotation. Once the absolute value of θ falls below a preset threshold, the roll correction process is completed.

Given the above information, the algorithm in Figure 29 was proposed to realize the concept of roll estimation. The existence of pitch and yaw made it difficult to calculate roll angles directly. Therefore, this algorithm useed an indirect approach in which the camera coordinate system was rotated step by step around the axis Z_c until the absolute value of θ fell below a threshold, which was 0.2° in this research; the opposite number of the total rotation angle was the roll angle. A bisection method was applied to approach a camera position without a roll angle. In this method, θ

was equally divided and measured repeatedly until its absolute value fell below the threshold. The flowchart in Figure 29 shows the details of the roll estimation algorithm.

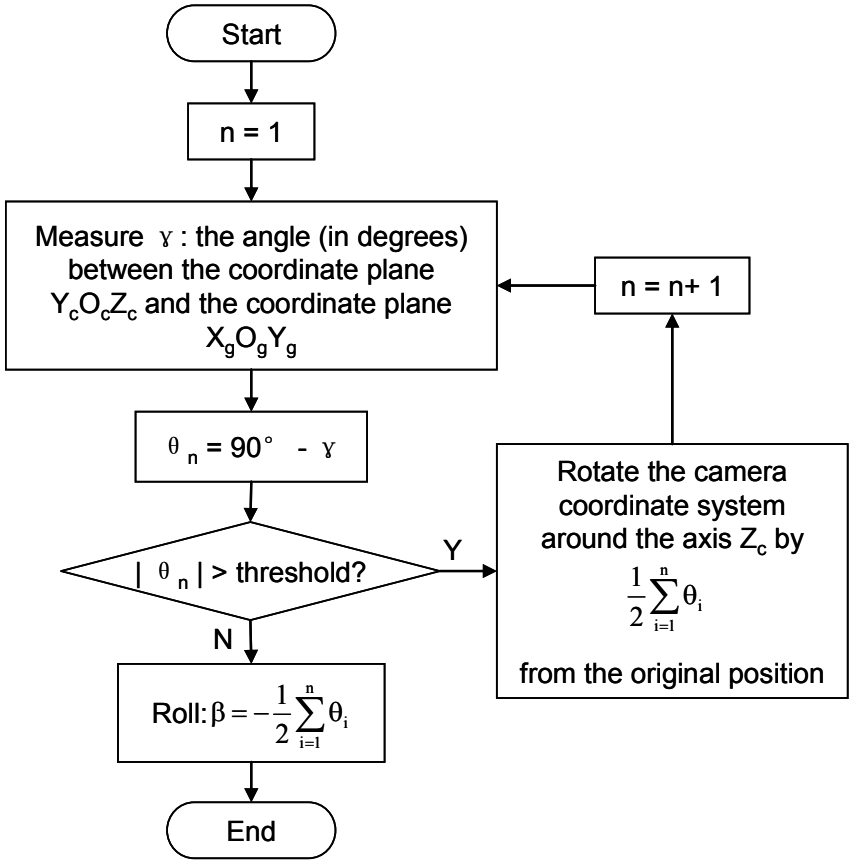


Figure 29. Roll estimation algorithm.

3.3.6 Camera Pitch Estimation

The camera pitch estimation was performed after completing the roll estimation. The pitch estimation began under an ideal camera position that had no roll. In reality, it started with a camera pose whose initial roll had already been compensated. As mentioned above, a typical camera installation pose for an agricultural vehicle involves the camera looking forward with a pitch angle pointing downwards at the ground. In this dissertation, the camera pitch estimation was designed to deal with this situation. The camera coordinate system was rotated around the axis X_c until the coordinate plane $X_cO_cZ_c$ was parallel to the coordinate plane $X_gO_gY_g$; the opposite number of the rotation angle was the pitch angle α (Figure 30). The angle between the axis Y_c and the axis Z_g,

named ε, was used to determine the rotation angle. Given that the axis Y_c and the axis Z_g are the norm vectors of two known planes $X_cO_cZ_c$ and $X_gO_gY_g$ respectively, ε was calculable. Because ε is supplementary to the rotation angle, the camera pitch angle can be represented as $\alpha = -(180° - \varepsilon)$.

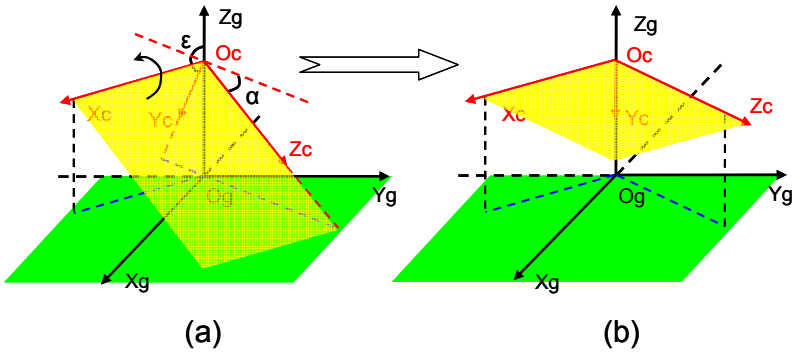


Figure 30. Camera pitch estimation.

3.3.7 Camera Yaw Estimation

The second task in the camera pose auto-calibration method was to determine the vehicle motion direction in the camera coordinate system for yaw estimation. Using the 3D information provided by the stereo camera, the tracked motion in the 2D image space of Figure 31 was transformed into the 3D camera coordinate system. By using the previous roll and pitch estimation results, the camera coordinate system was rotated to a new pose with only the yaw remaining, which was denoted as $X_{c1}Y_{c1}Z_{c1}$. Figure 32 shows the 3D feature motion vectors in $X_{c1}Y_{c1}Z_{c1}$. The variations of the vertical coordinates (Y_{c1}) of the vectors in $X_{c1}Y_{c1}Z_{c1}$ were within 100 mm, which meant that the ground features and their motions were transformed to a relatively level plane after the roll and pitch correction. The top view of the system $X_{c1}Y_{c1}Z_{c1}$ (Figure 33) shows that all of the motion vectors had almost the same orientation. Their angular deviations from the camera facing direction were caused by the camera yaw angle. Therefore, the average angular deviation of these motion vectors was calculated as the camera yaw estimation. For a more robust estimation, 15 consecutive image frames (about 2 m traveling distance of the vehicle) were applied to reduce the potential dynamic effects on the yaw estimation. The estimates from each two consecutive frames were averaged to reach the final yaw estimation.

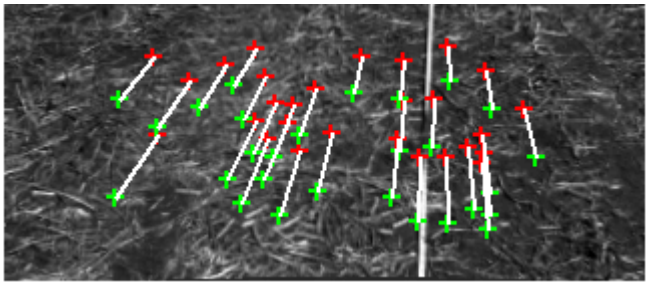


Figure 31. Ground feature tracking results between two consecutive image frames. The background picture is the first image frame. The red crosses and green crosses mark the feature positions in the 2D image coordinate system for the first frame and the second frame, respectively.

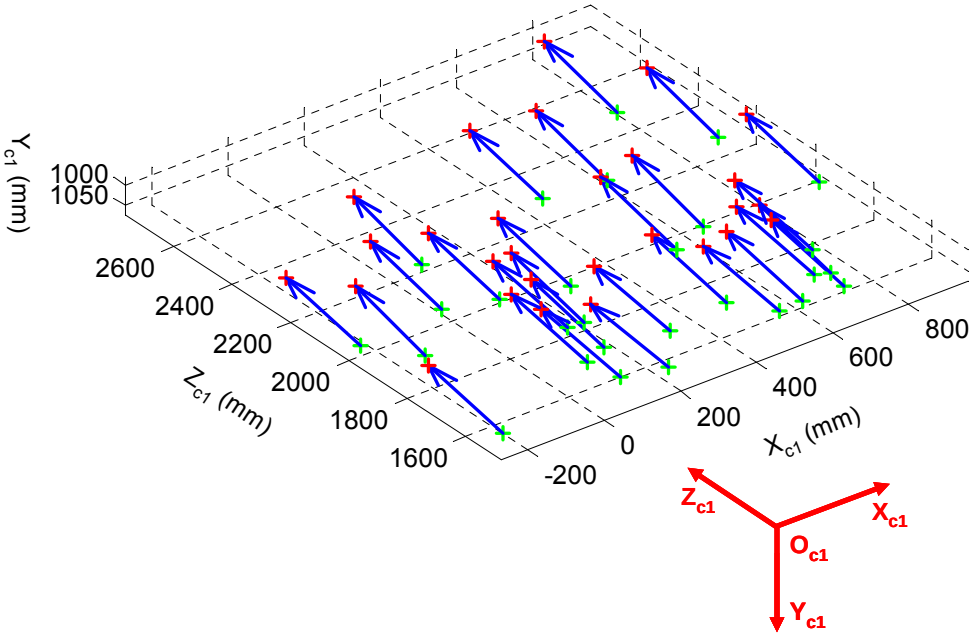


Figure 32. Feature motion vectors in the 3D space $X_{c1}Y_{c1}Z_{c1}$. $X_{c1}Y_{c1}Z_{c1}$ was obtained by correcting the roll and pitch for the camera coordinate system. All of the vectors point to the vehicle's heading direction.

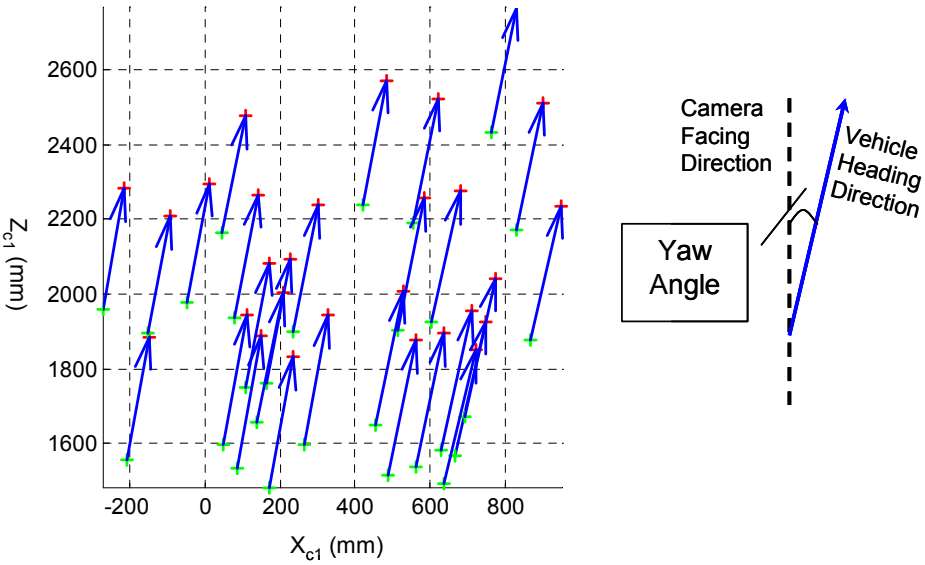


Figure 33. Top view of the 3D space $X_{c1}Y_{c1}Z_{c1}$.

3.4 DESIGN OF EXPERIMENTS

A series of field experiments were performed to validate the developed method for camera pose calibration. The design of experiments focused on four parts: (1) calibration field requirements; (2) camera pose settings; (3) pose estimation process; and (4) validation method for pose estimation.

3.4.1 Calibration Field Selection

The selection of an adequate field is based on two requirements for the proposed method: (a) the ground surface is relatively even, without significant ditches or bumps; (b) there are a sufficient number of visual corner features on the surface.

Four types of ground surfaces commonly found on US Midwestern farms were initially selected to evaluate their suitability for the camera pose auto-calibration. These ground surfaces were an unploughed soybean field, an unploughed corn field, a grass lawn and a gravel driveway outside of a garage (Figure 34). To satisfy the requirement of being relatively even, the selected fields had to be devoid of significant ditches or bumps. As regards the second requirement that there are

sufficient identifiable visual features, texture variety in terms of richness was tested in order to evaluate the applicability of the method. The lawn was the surface with the richest, but most uniform, texture, and the gravel driveway had the leanest, but most easily detectable, texture. The unploughed soybean field and the unploughed corn field fell between the lawn and the gravel driveway.

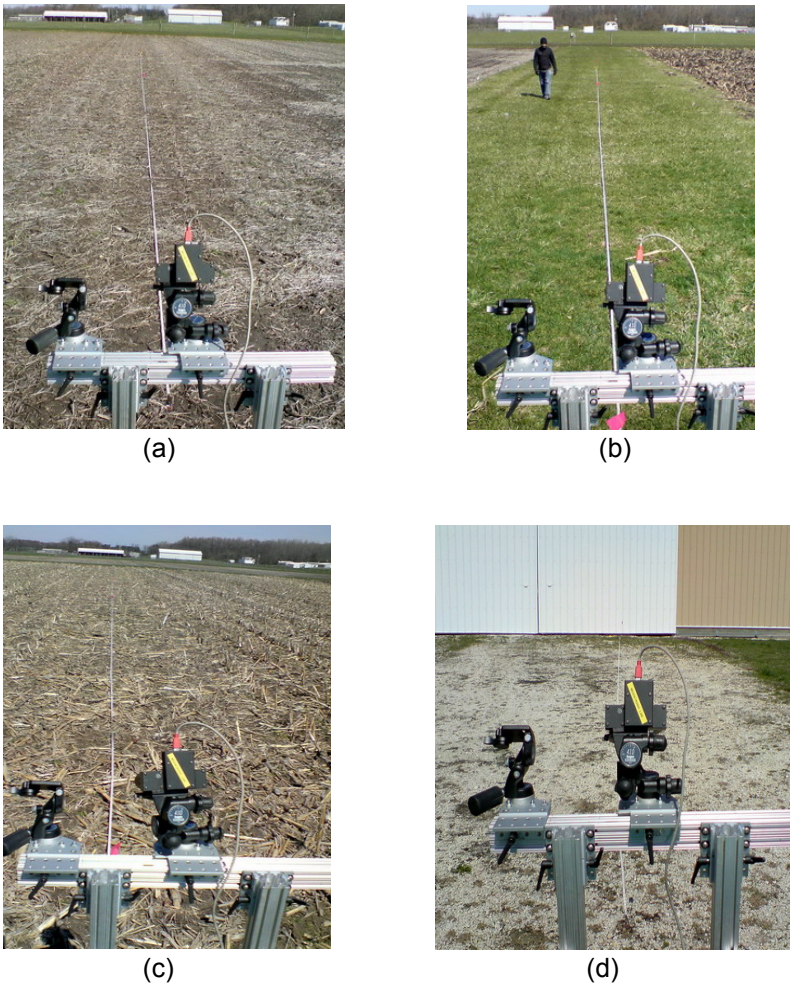


Figure 34. Types of ground surfaces commonly seen on US Midwestern farms. (a) Unploughed soybean fields, (b) lawn, (c) unploughed corn fields and (d) gravel driveways outside of garages.

3.4.2 Camera Pose Settings

Four settings for the camera pose were tested in order to simulate possible situations in real applications (Table 1). In setting 1, the camera was installed in a commonly used pose for vehicle

navigation, which meant that the camera looked forward with a pitch angle (about 25°) pointing downwards at the ground, without significant roll or yaw. Settings 2 to 4 were a series of tests concerning yaw variety. Setting 2 had the same configuration as in Setting 1. In setting 3, the camera yaw angle was increased by about 5° from the camera pose in setting 2. Another yaw increase of about 5° was implemented in setting 4 based on the pose from setting 3. The reason for the focus on camera yaw was that yaw has more influence than pitch and roll on vehicle navigation, particularly for finding the heading direction in a reliable manner. When the camera yaw angle was increased manually, the roll and pitch might change a small number of degrees, but the changes were visually insignificant.

Table 1. Camera pose settings for field tests.

Settings	Camera Pose		
	Roll	**Pitch**	**Yaw**
1	Insignificant	About 25°	Insignificant
2			
3	Insignificant	About 25°	Increase by about 5° based on setting 2
4	Insignificant	About 25°	Increase by about 5° based on setting 3

3.4.3 Pose Estimation Process

During calibration, the vehicle traveled slowly and straight forward (between 0.5 and 1 $m \cdot s^{-1}$) on relatively even ground for about 2 m. At the same time, the stereo camera recorded images continuously at a rate of four frames per second. The slow speed helped reduce vehicle vibrations that were caused by small bumps on the ground.

3.4.4 Validation Method for Pose Estimation

The pose estimation validation consisted of checking whether the calculated pose successfully reflected the relationship between the camera and the vehicle coordinate systems. As mentioned in the calibration algorithm section, static ground features were tracked in order to estimate the camera yaw. In other words, when the camera pose was transformed to a position without roll, pitch or yaw due to successful calibration, a repetition of the calibration test resulted in 0° being

the estimated yaw angle. This is the basic idea that was implemented in the validation algorithm.

In this study, a validation field that was over 10 m long was located next to a field about 2 m long that was used for calibrating the camera, as indicated in Figure 35. The vehicle acquired the images for the calibration and validation continuously during every test run. During validation, the vehicle also traveled straight forward and slowly, but traveled a distance that was longer than in the calibration test.

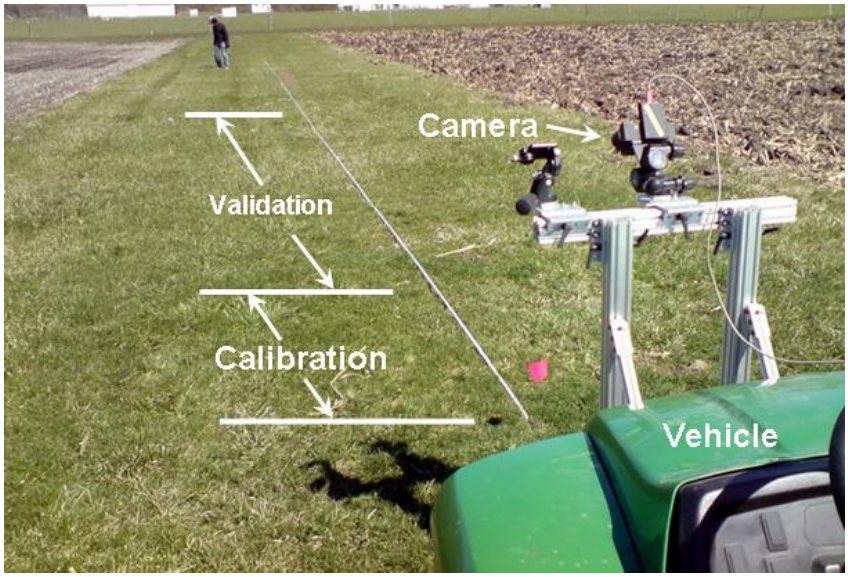


Figure 35. Field settings for camera pose calibration and its validation.

After the image acquisition phase was completed, the calibration algorithm was executed in order to process the images that were acquired in the field and calibrate the camera. After the pose was compensated, the developed validation algorithm used images that were acquired in the validation test in order to validate the pose compensation results. A flow chart of the validation algorithm is shown in Figure 36. The first two consecutive image frames taken from the validation field were fed into the algorithm. Using the same method as in the calibration algorithm, ground features were detected and tracked from this pair of consecutive images. For the purpose of validation, the estimated pose from the calibration test was imported into the algorithm to correct the current camera pose to a position that was devoid of roll, pitch or yaw. These corrected motions were then

used to estimate the camera yaw angle. If the camera pose calibration was correct, this camera yaw angle should have been close to 0°. A total of 40 pairs of consecutive image frames, covering about 10 m of traveling distance, were processed in the validation. If all of the calculated yaw angles from the 40 pairs were close to 0°, and the mean fell in an error range of ±1°, the pose calibration method was considered have been validated.

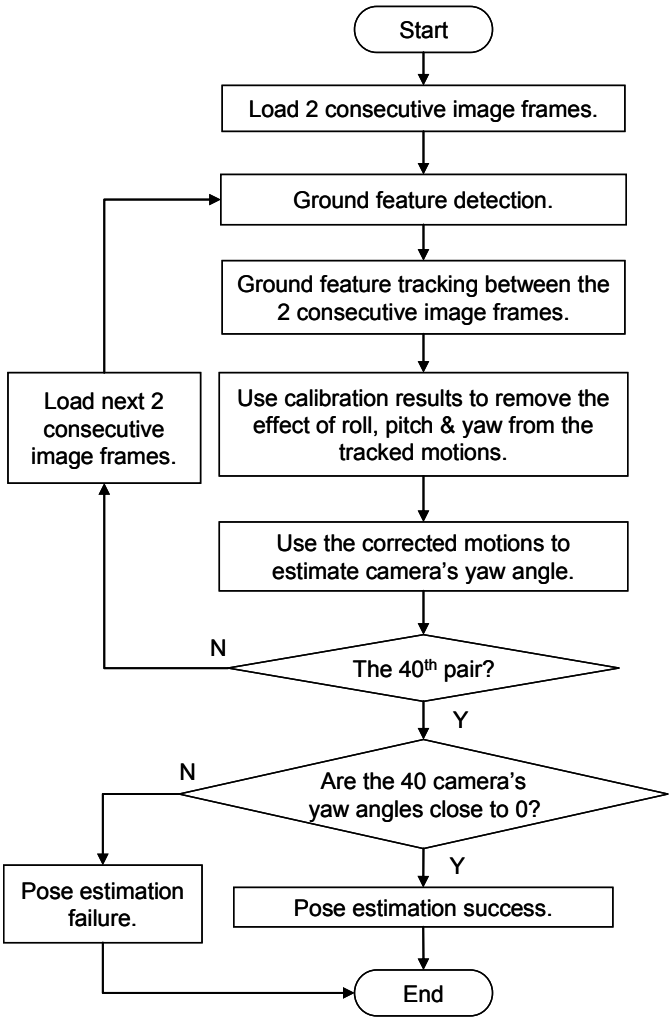


Figure 36. Validation algorithm for camera pose estimation.

3.5 FIELD VALIDATION TEST RESULTS

Field tests were conducted at the South Farms of the University of Illinois (Urbana, IL, USA) during the spring of 2009. In accordance with the experimental design, 16 runs, which consisted of

the four types of ground surfaces (Figure 34), were conducted in the field tests (Table 2). Each group of ground type contained four runs with different camera poses. During the first run of each group, the camera was installed according to pose setting 1 in Table 1. The second, third and fourth runs of each group used the settings 2, 3 and 4 of Table 1, respectively. The camera installation was conducted manually for each run, which could not sufficiently guarantee installation accuracy, therefore, two camera poses with the same setting, but from different runs, might be slightly different from each other; in addition, the pitch angles might deviate from 25° within a small number of degrees.

The calibration algorithm was executed first after the data acquisition stage. The estimated camera poses were obtained for all of the 16 runs except for the four lawn tests (Table 2). The estimated poses showed the changes that were induced by the specific camera settings. As regards pose setting 1, the absolute values of the estimated angles in runs1, 9 and 13 were around 5° for both roll and yaw, which complied with the designed pose with insignificant roll and yaw; the three estimated pitch angles for pose setting 1 were between 23° and 24°, which were close to the designed yaw angle (about 25°). As regards pose settings 2 to 4, all of the estimated yaw increases were within the range from 4° to 6° (Table 2), which successfully represented the designed yaw increase step (about 5°). Therefore, the calibration process worked on the unploughed soybean field, the unploughed corn field and the gravel driveway. The problem with the lawn experiments was that the calibration algorithm could not complete the ground feature tracking in a reliable manner.

Table 2. Field test results.

Run	Ground Type	Pose Setting	Camera Pose Estimation (degrees)				Camera Yaw Angles in Validation (degrees)	
			Roll	Pitch	Yaw	Yaw Increase	Mean	Standard Deviation
1	Soybean	1	-6.1	23.9	5.9	N/A	-0.1	0.7
2		2	-4.2	22.9	8.6	N/A	0.2	0.5
3		3	-4.0	24.3	12.9	4.3	-0.2	0.8
4		4	-7.5	24.9	17.2	4.3	0	0.8
5	Lawn	1	N/A					
6		2						
7		3						
8		4						
9	Corn	1	-6.6	23.0	4.4	N/A	0	1.2
10		2	-6.1	23.2	6.5	N/A	0.4	1.2
11		3	-5.6	26.1	11.2	4.7	0.4	1.2
12		4	-3.9	23.4	16.6	5.4	-0.4	1.5
13	Gravel Driveway	1	-4.9	23.8	5.8	N/A	-0.6	1.2
14		2	-6.5	24.5	5.4	N/A	0.9	1.6
15		3	-5.9	23.5	11.3	5.9	0.2	1.1
16		4	-4.1	22.5	17.2	5.9	-0.3	1.4

The validation algorithm was carried out after the pose estimation tests. Figure 37 shows the validation results for run 2 (soybean field, camera pose setting 2). After the camera pose was corrected using the estimated roll, pitch and yaw, the 40 camera yaw angles from the validation were close to 0°. Statistics showed that the average camera yaw angle found in the validation test was 0.2°, which fell within the expected error range of ±1°. Therefore, the proposed calibration method was validated for run 2. As regards the remainder of the tests, all of the runs except for the four lawn tests resulted in an average yaw angle within ±1° (Table 2). Therefore, the camera pose calibration method was validated for the unploughed soybean field, the unploughed corn field and the gravel driveway.

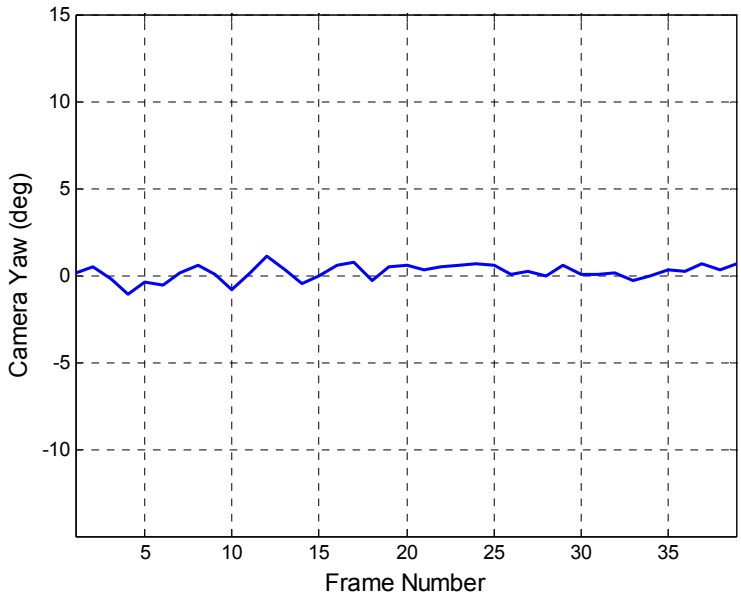


Figure 37. Camera yaw angles after roll, pitch and yaw correction.

3.6 DISCUSSION

Although the calibration method was validated for three types of farm ground, there remain two issues that should be examined in order to better evaluate the proposed method:

(1) How did the evenness of the calibration field affect the calibration?

(2) How did the difference in ground texture affect the calibration?

3.6.1 The Effect of Field Evenness

In the field validation tests that were performed, all of the selected fields were relatively even, though there was some unavoidable small bumpiness of the sort typically experienced in agricultural fields. However, the proposed calibration method was developed under the assumption that the ground surface was perfectly flat; this conflict resulted in fluctuations of the calculated yaw angle around 0° (Figure 37). The fluctuating curves of the yaw angle show an approximately periodic pattern, which suggests that the effect of bumpiness could be treated as white noise. Therefore, a moving average of period 5 was used to reduce the noise. Figure 38

shows the results of smoothing the data from Figure 37. Fluctuations were significantly reduced in the filtered plot.

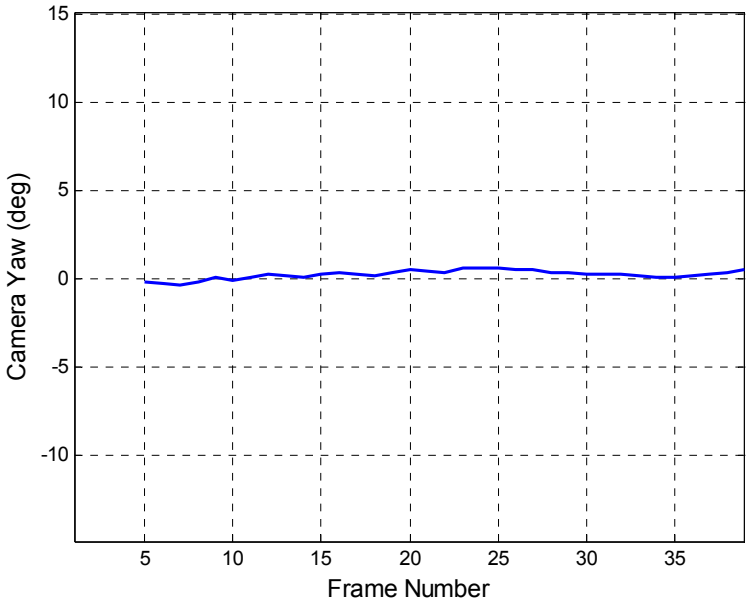


Figure 38. Camera yaw angles after applying a moving average of period 5.

The test results indicated that some significant bumpiness occurred during runs 13 and 14, which made the absolute values for the average yaw angle drift to the range of 0.5° to 1°, while all of the other runs stayed within the range of 0° to 0.5° (Table 2). The images taken for run 13 show significant vehicle (camera) rotations after frame 23, where the yaw angle curve starts to deviate from 0°, as represented in Figure 39. These significant rotations were caused by bumps or ditches in the test fields. As a result, irregularities should be avoided whenever possible when selecting a calibration field.

All of the test fields also exhibited noticeable differences in their evenness. As presented in Table 2, the test results obtained from the unploughed soybean field displayed smaller standard deviations for the estimated yaw (in the range of 0.5° to 0.8°) than did those from the unploughed corn field and the gravel driveway (in the range of 1.1° to 1.6°). The possible cause of this difference might be that the unploughed soybean field had less bumpiness than the other test fields.

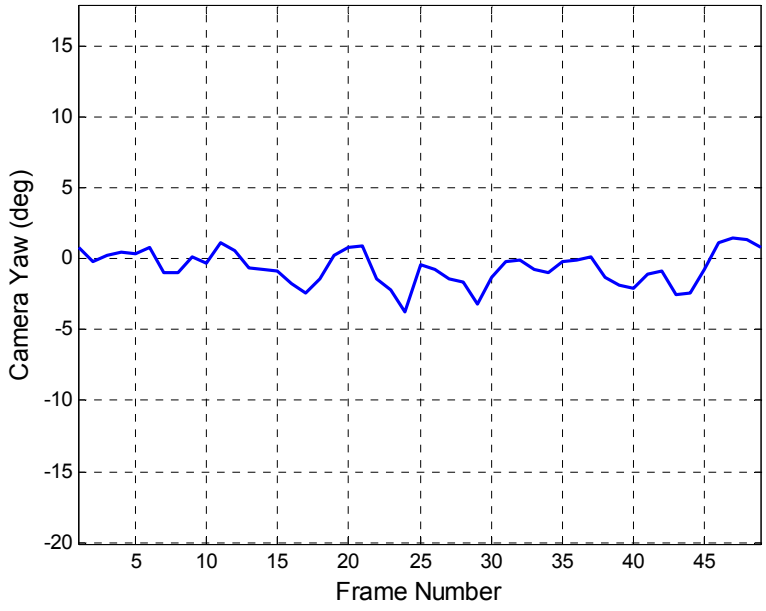


Figure 39. Camera yaw angles in the validation test for run 13.

3.6.2 The Effect of Ground Texture

Different ground textures can affect the calibration results in different ways. For example, comparing the typical texture features on the four types of test ground (Figure 34) revealed that lawn texture was more homogeneous (had fewer patterns) than were the others, which made it difficult to identify and track ground features in a reliable manner. Consequently, the calibration method failed on the lawn because the detected ground features were too uniform. Therefore, when selecting calibration fields, ground with homogenous textures should be avoided.

3.7 CONCLUSIONS

An automatic calibration method was developed for a stereo camera that was installed on an agricultural vehicle in order to estimate its installation pose. This method used a set of consecutive image frames captured by a vehicle-mounted stereo camera to detect and track natural texture features on typical agricultural grounds, and then calibrated the camera installation pose. A validation method was also developed to test the proposed calibration algorithm. Field tests

showed that the calibration method could successfully estimate camera poses in the environments of relatively even agricultural fields and gravel driveways.

CHAPTER 4: A METHOD OF ESTIMATING HEADING DIRECTION FOR STEREOVISION-BASED VEHICLE NAVIGATION IN OPEN AGRICULTURAL FIELDS

4.1 INTRODUCTION

Machine vision has been the subject of research for the auto-guidance of farming vehicles since the mid-1980s. One major application is the use of machine vision to imitate human visual perception in order to detect crop rows, and the use of rows to determine vehicle guidance direction in farming operations. The literature review in Chapter 2 shows that most of the reported research has dealt with fields that feature crop rows or cut and uncut edges (structured environments); however, some farming operations, such as ploughing and seeding, occur in open fields (ill-structured environments). The difference with respect to vision-based navigation in the two cases is that there are few significant visual clues available in open fields that can provide the necessary feature points for finding a guidance direction. Few researchers have used machine vision for such applications. Leemans and Destain (2006) used the Hough transform to detect furrows made by a planter in previous drills, and then determined the guidance direction for a seeding vehicle. The method depends on manmade furrows to provide structures in the environment, and remains a good research exploration in the area of machine vision-based automatic guidance for ill-structured environments. Two issues concerning this method may increase interest in future research:

(1) This method works for the majority of planting operations, but not for the first drill of planting, because there are no existing furrows that can be used as references.

(2) Machine vision may not be as flexible as human vision, but is usually capable of repeating an algorithm rapidly and reliably for long periods of time. Therefore, it would be interesting to determine whether it is possible to find new guidance references that may be difficult for human drivers to track while being considerably easier for machine vision.

This dissertation was inspired by these two open issues, and investigated the feasibility of using

machine vision to determine an agricultural vehicle's heading in an open field without manmade references. By way of preliminary exploration, the proposed algorithm applied an off-line process to the acquired images and dealt only with straight forward navigation (keeping a constant heading direction), which is a normal working mode for planting or ploughing. Small turns (such as oscillations around a desired heading direction) were allowed, but sharp turns or U turns were not within the scope of this research. Therefore, the specific objectives of this research project were:

(1) Use machine vision to detect static natural features in an open field.

(2) Investigate whether it is possible to use the detected natural features as references that can be used to determine the vehicle heading direction when the vehicle travels in straight lines.

4.2 THE CONCEPT OF VEHICLE HEADING DIRECTION ESTIMATION USING STEREOVISION.

In accordance with the objectives, two concerns appeared during the development of the method for determining the heading of a vehicle:

(1) How to select appropriate static natural field features that are detectable by machine vision.

(2) How to use those features as reference points for the purpose of estimating the heading direction of the vehicle.

4.2.1 Static Natural Field Feature

An open field usually has an ill-structured appearance, which makes it difficult for human drivers to find reference points that can be used to steer their vehicles. However, there are some static natural features that remain in the field. As shown in Figure 40, the crop residue, which is normally seen in open fields, differs from the soil in color; on the other hand, residue simply lies in the field randomly. Therefore, although it is difficult to detect any special patterns for the entire field, the local appearance of a field with residue is unique and can be distinguished. For example, window 1

and window 2 in Figure 40 are two local views of an open field that have different textures that can be easily distinguished by the human eye. Machine vision is also capable of detecting such local features (Li and Allinson, 2008). Therefore, the selection of appropriate static natural features can be realized using machine vision.

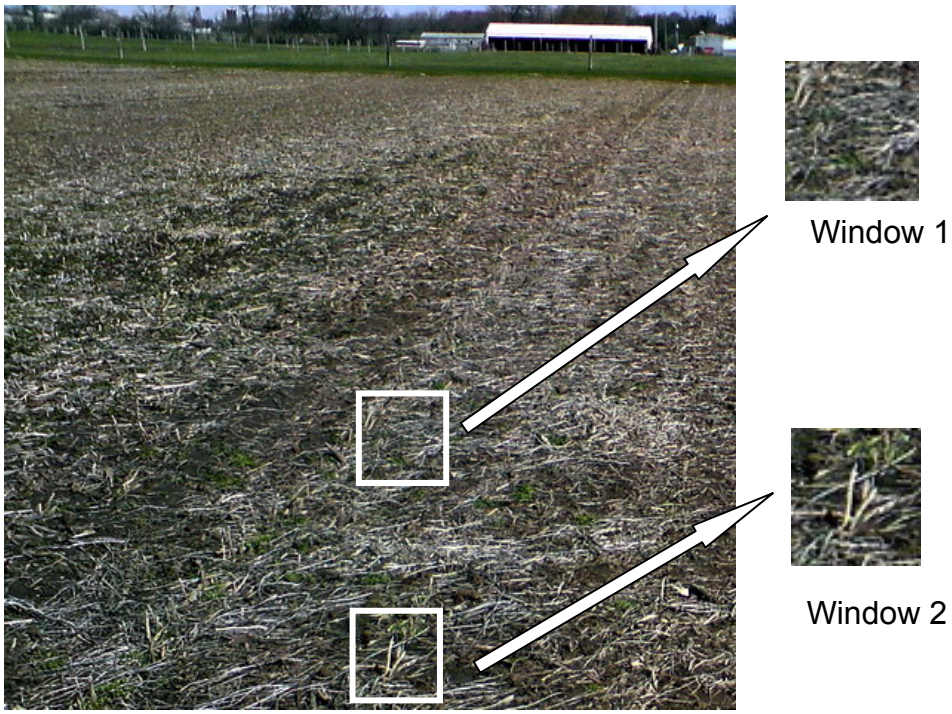


Figure 40. An ill-structured open field.

4.2.2 Reference for Heading Direction Estimation

The selected static field features can become reference points for describing the motion of a vehicle if they are trackable through the use of a camera that is mounted on the vehicle. Several studies have shown that machine vision can track local natural features in off-road environments (Cheng et al., 2006; Nister et al., 2006; Agrawal et al., 2007). Note that an open agricultural field is a typical off-road environment, and detected natural features are trackable using machine vision. If stereovision is used, the relative motion of the vehicle with respect to reference points in 3D space is also available. The 3D motion of the vehicle is an indication of its heading direction.

4.2.3 Concept Description

The above information was essential for the elaboration of the concept of using machine vision for the purpose of estimating vehicle heading direction. A stereo camera mounted on a vehicle was used for image acquisition. Figure 41 shows the block diagram of this concept. First, natural features were detected in two consecutive 2D images that were taken by the stereo camera when the vehicle was in motion; second, the features were tracked across the two images in 2D space; third, the depth information provided by the stereo camera was used to reconstruct the tracking results in the 3D space of the open field; fourth, the relative motions of the static features with respect to the vehicle were used to obtain the vehicle motion in the open field during the time that elapsed between two image frames. At the end, the heading direction was calculated using the estimated vehicle motion.

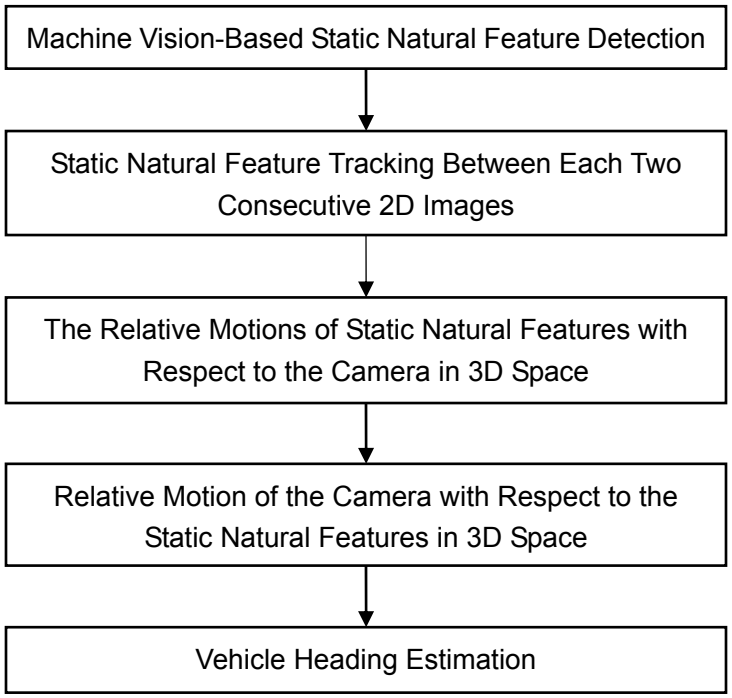


Figure 41. Vehicle heading estimation with stereovision: block diagram.

4.3 ALGORITHM DEVELOPMENT

A machine vision algorithm was developed in order to realize the proposed concept of vehicle heading estimation. The details of the algorithm are provided in this section.

4.3.1 System Architecture

The experimental vehicle used in this research was a John Deere Gator™ Utility Vehicle (Figure 11). A binocular stereo camera manufactured by Videre Design (Menlo Park, California, USA) was mounted on a rigid frame in the front part of the vehicle. The camera looked forward with a pitch angle aimed downward at the ground. An onboard computer was used to store the images that were acquired by the camera.

4.3.2 Definition of Coordinate Systems

Before discussing the heading estimation algorithm, it is necessary to define the coordinate systems. The coordinate systems were defined as being the same as in Chapter 3, the details appear in section 3.2.2.

4.3.3 Static Natural Feature Detection and Tracking

Static natural feature detection and tracking in 2D image space used the same algorithm that was described in sections 3.3.2 and 3.3.3. An example of the 2D feature tracking results is shown in Figure 42.

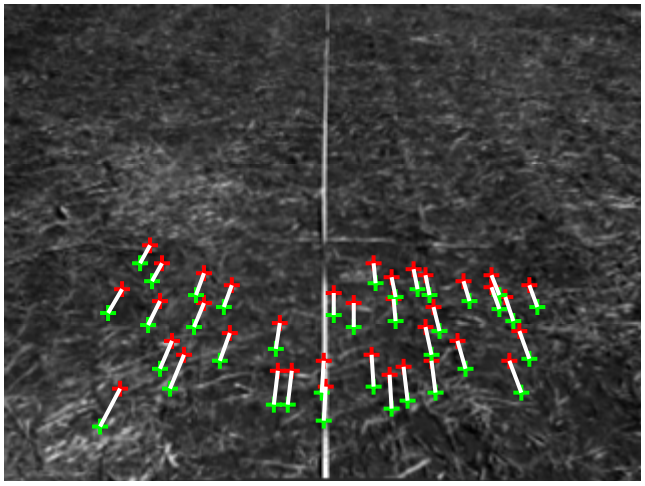


Figure 42. Ground feature tracking results between two consecutive image frames after removing outliers using a two-stage RANSAC process. The background picture is the first image frame.

All of the tracked motion vectors had to be transformed from the 2D image space to the 3D ground coordinate system, in order that the vehicle motion in the field could be described. As mentioned in the definition of the coordinate systems, the ground coordinate system was defined at the starting point of a navigational path. In the feature tracking process across two consecutive image frames, the starting point is the place where the vision system takes the first image frame. Using the depth information provided by the stereo camera and the geometrical relationship between the camera coordinate system and the ground coordinate system (obtained from the extrinsic parameters of the camera calibration), allows the tracked motions to be converted to the ground coordinate system. Figure 43 shows the transformation results for all of the tracked 2D points in Figure 42. The beginning and ending crosses of the direction arrows mark the feature positions from the second image and the first image respectively. The relative 3D motions of these static reference points in the ground coordinate system were caused by the motions of the vehicle. Therefore, the motion vectors pointing from the second positions to the first positions of the features show the relative motion of the vehicle with respect to the reference points.

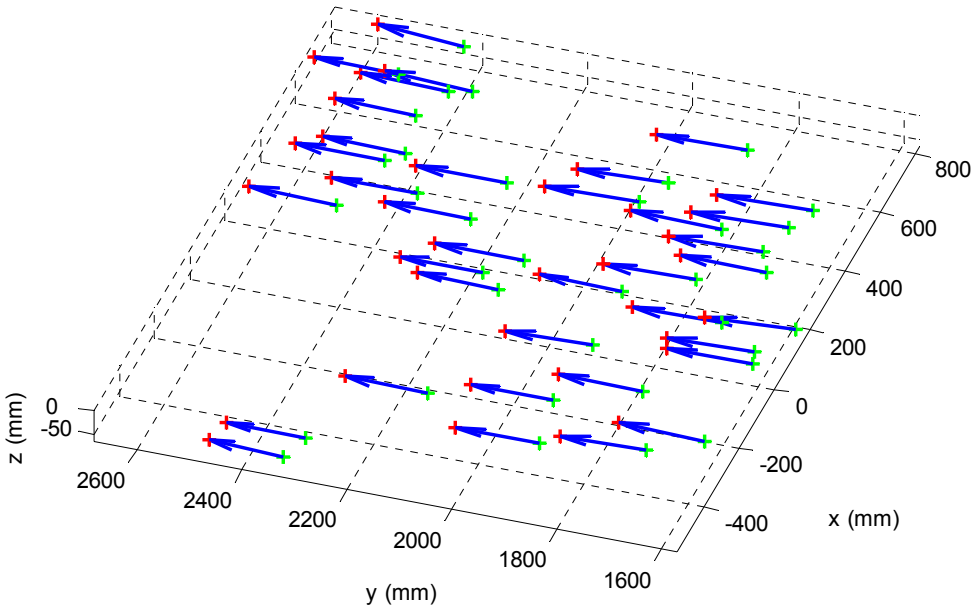


Figure 43. Feature tracking results in the 3D ground coordinate system.

4.3.4 Heading Direction Estimation

In this research, the vehicle heading direction was defined as shown in Figure 44. A heading angle $\lambda = \arctan \Delta x/\Delta y$ was used to represent the heading direction, where Δx and Δy were the vehicle movements on the x and y axes respectively in the ground coordinate system. The angle ranged from -90° to 90°; when the vehicle moved straight forward, $\lambda \approx 0°$; when the vehicle turned to the right side, $\lambda > 0°$, and vice versa. According to the objectives of this research, no significant rotation was expected to occur during the movements of the vehicle; therefore, it was assumed that there were only translational movements for the vehicle and the mounted camera.

A top view of the feature tracking results is shown in Figure 45. All of the motion vectors point in almost the same direction, which is the heading direction of the vehicle. The heading angles for all of the motion vectors were calculated, and their mean value was used as the final heading direction.

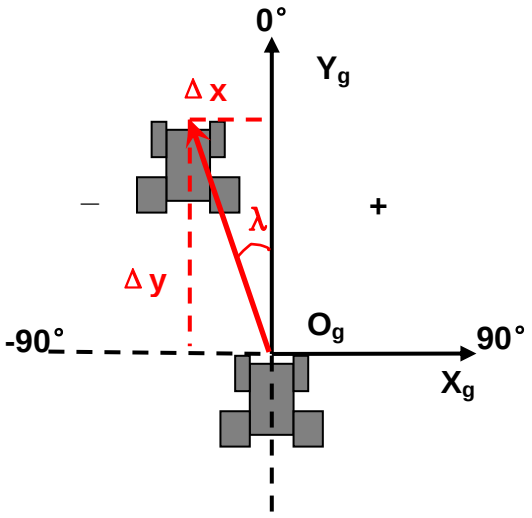


Figure 44. The definition of the vehicle heading direction.

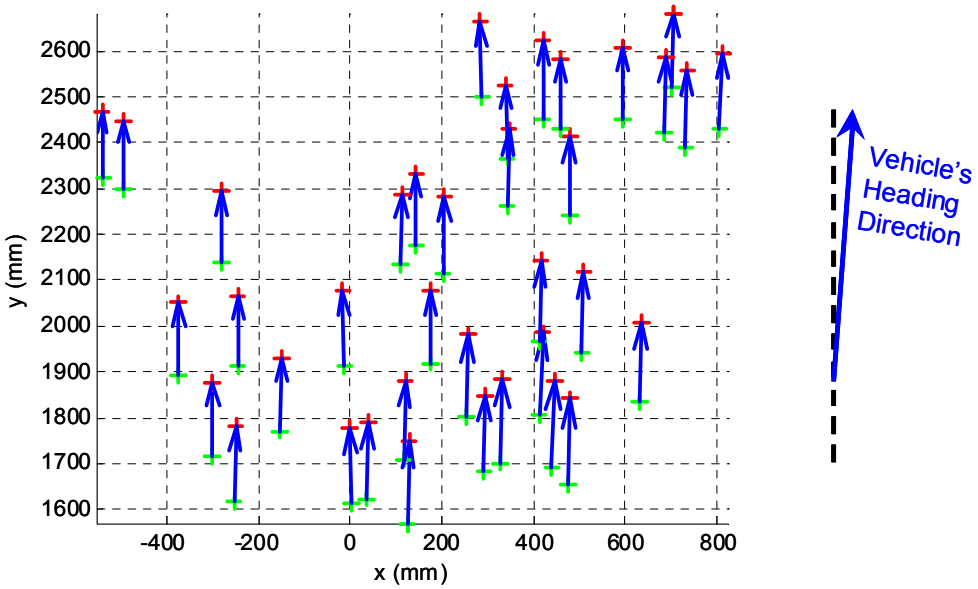


Figure 45. The top view of the motion vectors in the ground coordinate system.

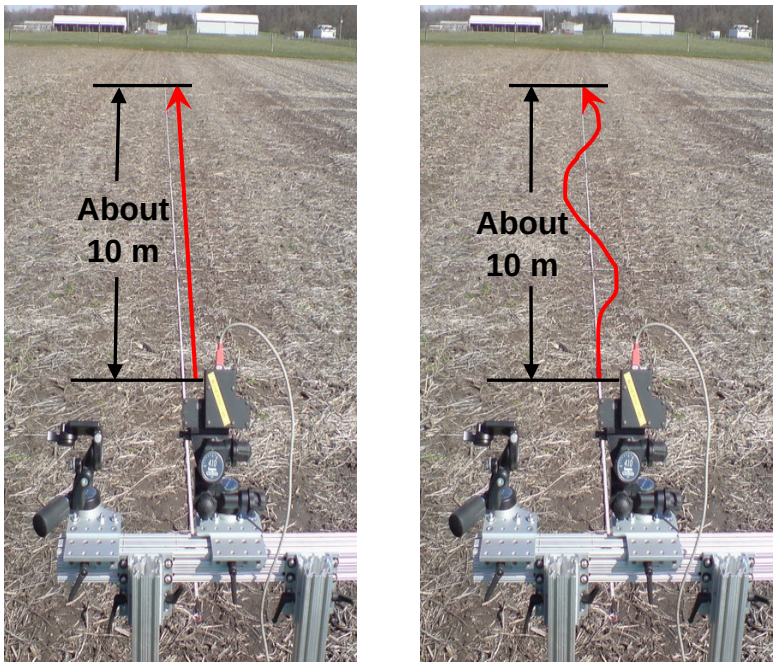
4.4 DESIGN OF EXPERIMENTS

4.4.1 Experiment Process

This research focused on determining a vehicle heading when the vehicle traveled in an open field

without taking sharp turns. Complying with this objective required two kinds of traveling modes and two kinds of open fields for field tests.

The two traveling modes were straight mode and oscillating mode. As shown in Figure 46, a white straight tape was laid in an open field, and this was used as a reference by the driver in order to steer the vehicle. In the straight mode (Figure 46(a)), the vehicle traveled along the straight tape for about 10 m; in the oscillating mode (Figure 46(b)), the vehicle oscillated around the straight tape while traveling for about 10 m alongside the tape.



(a) Straight Mode (b) Oscillating Mode

Figure 46. Vehicle traveling modes tested in the experiments.

The two kinds of open fields were both untilled but used different crop residues. One had soybean residue, while the other one had corn residue. Due to the different shapes of the residue, the two fields differed in appearance.

There were four combinations of the two traveling modes and the two kinds of open fields (Table 3). In order to test the reliability of the proposed algorithm, each combination was tested for

three runs.

Table 3. Four groups of experiment settings.

Combination	Run	Traveling Mode	Residue Type
1	1-3	Straight	Soybean
2	4-6	Straight	Corn
3	7-9	Oscillating	Soybean
4	9-12	Oscillating	Corn

The vehicle traveling speed was not exactly constant, but was tried to keep at about 0.7 $m \cdot s^{-1}$ in the field tests. This speed was slower than that in normal seeding or ploughing operations, but was able to fulfill the purpose of feasibility investigation of the concept for heading estimation. While traveling, the stereovision system shot and recorded image frames continuously at a rate of four frames per second. The recorded images were post processed in order to estimate the heading directions of the vehicle.

4.4.2 Evaluation Method

One obstacle to the evaluation of the field tests was that high accuracy global positioning systems (GPS) and inertial measurement units (IMU) were not available due to unexpected technical failures of those systems. Therefore, no position information was recorded as a reference for conducting a quantitative evaluation. However, because the research goal was to investigate the feasibility of the proposed concept, based on the experiment process, there were alternative methods of evaluating the usability of the proposed concept. Given that the heading direction was calculated using the change of the vehicle position, if the estimated position change was correct, the heading direction would be correct. This idea was used to develop the following evaluation method.

The vehicle was driven along the straight tape in either straight mode or oscillating mode during the testes; therefore, the straight tape always showed in the view of the camera for both modes. Since the straight tape was stationary, it can be used as a reference to calculate the motion of the vehicle in the left and right direction, which was the X_g axis in the ground coordinate system. An example from a test run is shown in Figure 47, which contains three image frames that were

selected from a serious of images taken when the vehicle was in motion. The positions of the straight tape in the three images were different, which was caused by the oscillating motion of the vehicle around the straight tape. This oscillating motion was slight but did exist in the straight mode tests due to the inconsistence of manual steering; while, it was significant in the oscillating mode tests. The intersection of the straight tape and row 220 of an image (the dash line in each frame of Figure 47) was used to determine the oscillation in a quantitative manner. The coordinates of the intersection were measured manually from each image as shown in Figure 47. The position changes of the intersection in horizontal direction between every two consecutive images were caused by the motion of the vehicle along X_g axis in the ground coordinate system. Therefore, the transformation results of these intersection coordinates from the 2D image space to the 3D ground coordinate system represented the motion of the vehicle along X_g axis, but with opposite directions. The error of the manual positioning of the intersection in a 2D image was no more than 1 pixel, which corresponded to about 5 mm after the 3D transformation to the ground coordinate system. Therefore, this measurement for the motion of the vehicle along X_g axis was relatively accurate, and the results were used as a benchmark in evaluating the vehicle motion estimation along X_g axis. As regards the vehicle motion along Y_g axis (the forward direction) in the ground coordinate system, there was not a proper method that was found to provide a relatively accurate measurement based on the current experimental conditions.

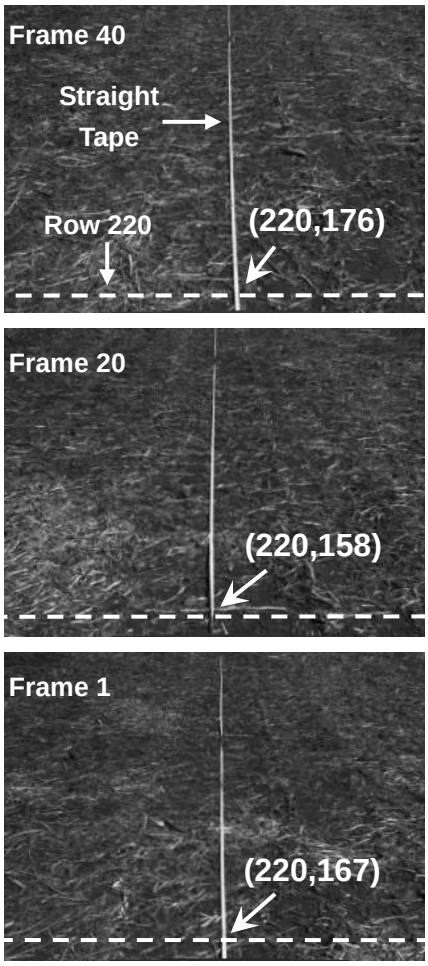


Figure 47. The positions of the straight tape in three image frames that were selected from a series of images taken continuously when the vehicle was in motion.

The position changes of the vehicle between each two consecutive image frames were estimated by the proposed heading estimation method. Using the concept of dead reckoning, all of the position changes were connected and the vehicle trajectory was reconstructed in the field. An example is shown in Figure 48, where the vehicle started at point O_g, which was the origin of the ground coordinate system. The position change (arrow D1), which was estimated from image frames 1 and 2, took the vehicle to a new position represented by the coordinate system $X_{g1}O_{g1}Y_{g1}$. The arrow D2 obtained from image frames 2 and 3 was the position change from O_{g1} to O_{g2}. Finally, the vehicle reached the position O_{g3}. Given that the vehicle was assumed to have only translational movements, all of the four coordinate systems ($X_gO_gY_g$, $X_{g1}O_{g1}Y_{g1}$, $X_{g2}O_{g2}Y_{g2}$ and $X_{g3}O_{g3}Y_{g3}$)

had parallel axes. Therefore, the connection of D1, D2 and D3 in the ground coordinate system was an approximation of the vehicle trajectory. This estimation result was compared with the manually measured vehicle positions to check the deviation of the estimated trajectory from the real trajectory. As mentioned above, the measured vehicle motion did not contain the motion along Y_g axis, which impeded a complete comparison between the estimated trajectory and the real trajectory. However, the deviation along X_g axis after the vehicle traveling for about 10 m was still an indicator for evaluating the position estimation. If the position estimation was accurate, the deviation should be close to 0. This evaluation method was valid for the purpose of feasibility investigation in the dissertation, and was applied to the images acquired from the field tests.

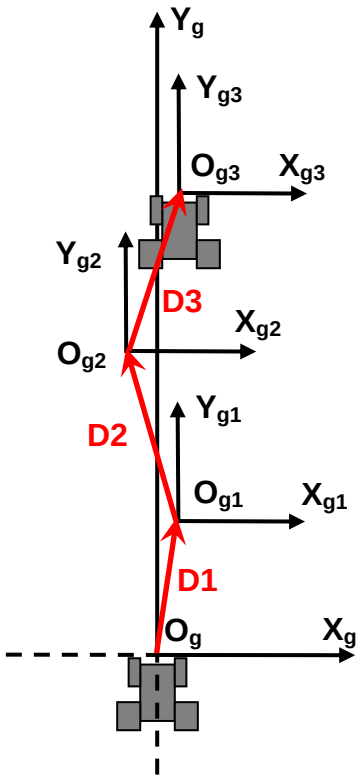


Figure 48. Reconstruction of the vehicle trajectory by connecting the position changes between each two consecutive image frames.

4.5 RESULTS

Field tests were conducted at the South Farms of the University of Illinois (Urbana, IL, USA)

during the spring of 2009. The results are presented in this section.

4.5.1 Straight Mode

The heading estimation and its evaluation were conducted in runs 1-6 with straight traveling mode. Figure 49 is an example of trajectory reconstruction from run 3. The estimated trajectory was fairly straight throughout the entire traveling distance, which was a proper reflection of the actual motion of the vehicle. The deviation of the estimated vehicle position from the manually measured vehicle position along X_g axis is shown in Figure 50. The mean of the deviations was 5 mm, and the maximum deviation (absolute value) was 14 mm. These two values showed that the estimated vehicle positions were fairly close the to the real vehicle positions in run 3. Table 4 summarizes the results for straight mode. As regards all 6 runs expect run 2, the deviation means were between -27 mm and 5 mm; the maximum deviations (absolute value) were no more than 48 mm. Compared with the total traveling distance, which was around 10 m, the means and maximum deviations were fairly small in the 5 runs. Therefore, the estimated vehicle positions were proved to be close to the real positions in these 5 runs with straight mode. Run 2 resulted in a mean of -60 mm and a maximum of 109 mm, both of which showed higher estimation errors than other straight mode runs. Figure 51 demonstrates the results from run 2, in which the deviation of the estimated positions from the real positions grew as the increase of the traveling distance. The divergence between these two curves might be caused by an improper camera installation pose calibration, which will be discussed in the discussion part of this chapter.

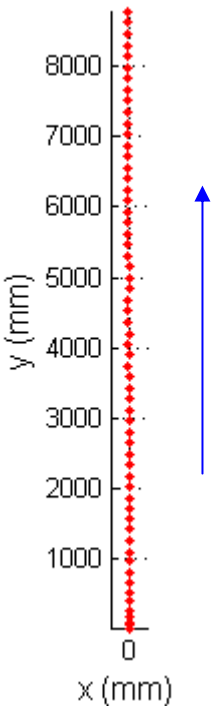


Figure 49. Estimated trajectory of run 3.

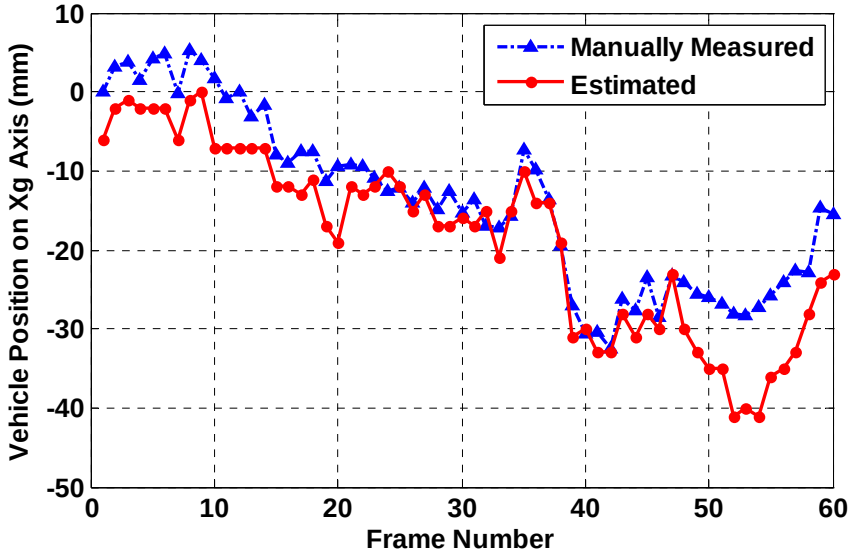


Figure 50. The estimated and manually measured vehicle positions on X_g axis from run 3.

Table 4. Results of the straight mode runs.

Run	Deviation (mm)	
	Mean	Maximum (Absolute Value)
1	-13	48
2	-60	109
3	5	14
4	-27	47
5	-4	22
6	-14	39

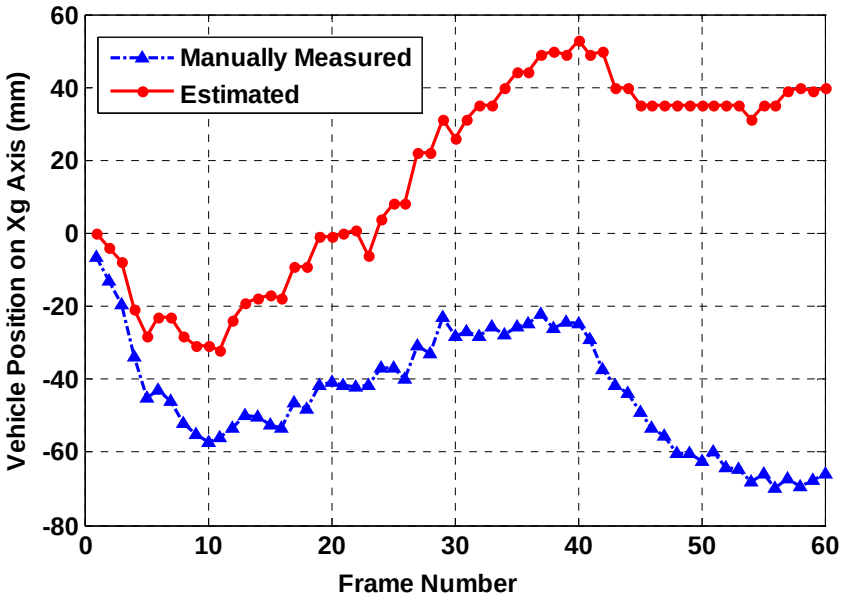


Figure 51. The estimated and manually measured vehicle positions on X_g axis from run 2.

4.5.2 Oscillating Mode

During the oscillating mode tests (runs 7-12), the estimated positions reflected the oscillating motions properly for all of the runs. For example, Figure 52 shows the deviation of the estimated vehicle position from the manually measured vehicle position along X_g axis. The mean of the deviations was -13 mm, and the maximum deviation (absolute value) was 53 mm. Therefore, the estimated vehicle positions were fairly close to the real vehicle positions in run 8. Table 5 summarizes the results for oscillating mode. As regards all 6 runs, the deviation means were between -28 mm and 48 mm; the maximum deviations (absolute value) were no more than 132

mm. The runs with oscillation mode showed more errors than the runs with straight mode in vehicle position estimation along X_g axis. This difference showed that the proposed method had limitations in dealing with vehicle turns, which will be discussed in the discussion part of this chapter. However, compared with the total traveling distance, which was around 10 m, the means and maximum deviations were fairly small in the 6 runs. Therefore, the estimated vehicle positions were proved to be fairly close to the real positions in these 6 runs with straight mode.

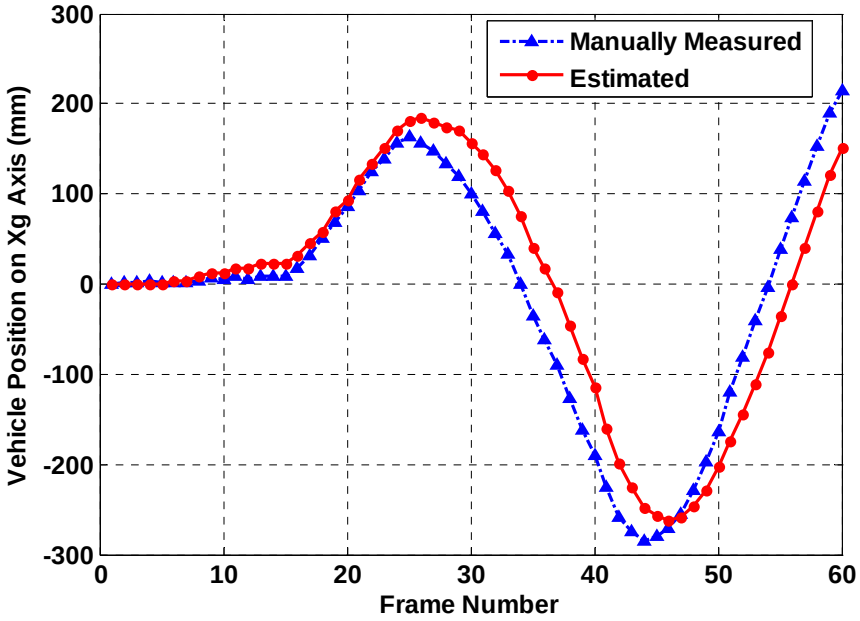


Figure 52. The estimated and manually measured vehicle positions on X_g axis from run 8.

Table 5. Results of the oscillating mode runs.

Run	Deviation (mm)	
	Mean	Maximum (Absolute Value)
7	48	132
8	-13	53
9	-21	69
10	-28	100
11	-16	75
12	-17	63

4.5.3 Summary

The developed algorithm worked accurately in estimating the vehicle position along X_g axis in 5 of the 6 runs with straight mode; the estimation accuracy degraded in the runs with oscillating mode, but was still acceptable. No significant difference was noticed between the results from the soybean field tests and the corn field tests. Therefore, the algorithm for vehicle heading estimation was partly validated, but a complete validation required an accurate measurement of the vehicle motion along Y_g axis (the forward direction) in the ground coordinate system, which was not able to be realized during the existing field tests. The field test results also indicated that the camera installation pose calibration and the vehicle turns might be the two issues that affected the vehicle heading estimation accuracy.

4.6 DISCUSSION

4.6.1 Camera Installation Pose Calibration

The proposed algorithm transformed the tracked motions from the 2D image space to the 3D field space, so it was necessary to know the geometrical relationship between the camera coordinate system and ground coordinate system. It was assumed that the vehicle had only translational movements, and on the other hand, an open field is usually relatively even, so this geometrical relationship was considered to be constant for the estimations in all of the image pairs. Therefore, a high quality calibration of this geometrical relationship was needed; otherwise, the heading

direction error would affect the vehicle motion estimation continuously and lead to navigation failure. An example is given in Figure 53; if a poorly conducted calibration results in an error of 1° for the heading estimation, a deviation of 0.17 m in the estimated trajectory would occur after traveling 10 m in a straight line. The proposed automatic method for camera pose calibration in Chapter 3 was used to determine the geometrical relationship between the camera and the ground coordinates. The results for straight mode (Table 4) showed that the absolute values of the deviation means in all the 6 runs except run 2 were less than 30 mm, which proved that the calibration method provided an acceptable camera installation pose estimate. However, run 2 showed a divergent estimation for the vehicle positions on X_g axis in the ground coordinate system (Figure 51), and resulted in a maximum deviation of 109 mm over the traveling distance. Therefore, a camera installation calibration error should exist in run 2 and cause this divergence.

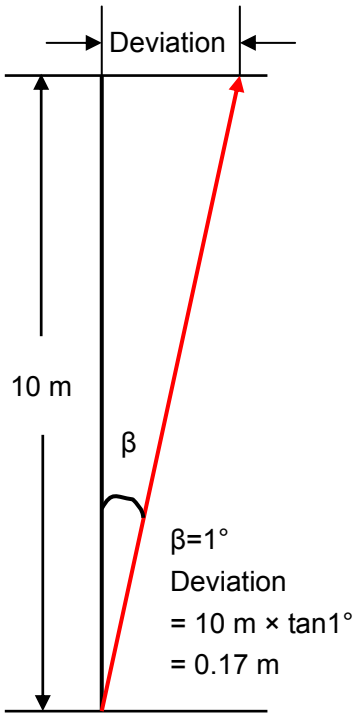


Figure 53. An example of how heading estimation errors affect navigation.

4.6.2 The Limitation From the "No Rotation" Assumption

The proposed algorithm was based on the assumption that there were no significant camera rotations while the vehicle was in motion. Two parts of the algorithm relied on this assumption, one was in stage 2 when using RANSAC to remove outliers, and the other one was in the heading angle calculation. Therefore, these two parts will not work anymore if significant camera rotations occur and the motion vectors do not exhibit a homogenous orientation. This limitation provides the reason why the estimation results from the runs with oscillating mode were inferior to those from the runs with straight mode. Slight camera rotations were introduced to the tests by the oscillating motion of the vehicle, which violated the assumption.

This assumption is valid for a feasibility investigation for the concept of heading estimation. However, in an off-road environment, ground bumpiness and vehicle turns can both produce camera rotations. In order to improve the estimation accuracy and to make this algorithm more applicable, having the capability of dealing with camera rotations is recommended for future research.

4.6.3 Real-Time Issues

The summarized work was intended for a feasibility investigation, so a real-time application was not included. However, this may become a major concern in the future. If optimized codes and fast processors are used, it is possible to realize real-time processing of this algorithm.

4.7 CONCLUSIONS

An image processing algorithm was developed to investigate the feasibility of using stereovision for the purpose of estimating the heading direction of a moving vehicle in open agricultural fields. The algorithm detected static natural ground features in corner shape, and used them as reference points when estimating vehicle heading directions. Field tests were conducted for the purpose of

algorithm evaluation. The algorithm worked accurately when the vehicle traveled straight forward. When the vehicle traveled in an oscillating mode, the algorithm was able to reflect the heading direction of the vehicle; however, due the existence of slight camera rotations, the heading estimation performance in the oscillating traveling mode was inferior to that in a straight traveling mode. In summary, the proposed algorithm and the field tests showed that it is possible to use stereovision to estimate the heading direction of a moving vehicle in open agricultural fields.

CHAPTER 5: FIELD EDGE DETECTION USING COLOR SEGMENTATION IN OPEN FIELD ENVIRONMENTS

5.1 INTRODUCTION

Machine vision applications for farming vehicle navigation can be divided into two categories: the first is applications in structured environments such as fields with crop rows; the second is applications in ill-structured environments such as open fields without crops. Machine vision was introduced to agricultural automatic guidance in the mid-1980s, and since then, most of the research has focused on applications in structured environments (Reid and Searcy, 1987; Marchant, 1996; Billingsley and Schoenfisch, 1995; Ollis and Stentz, 1997; Pinto and Reid, 1998; Benson et al., 2000(a); Han et al., 2004; Kise et al., 2005; Rovira-Más et al., 2005). However, due to the lack of significant visual clues, few of the research studies have developed applications for use in ill-structured environments. A recent research study conducted by Leemans and Destain (2006) explored this particular application. They used a camera to guide a planting vehicle by following the furrows that had been made by a planter during its previous drill. This might work for the majority of planting operations, but not for the first drill of planting because there are no furrows that can be used as references in that case. In non-computer assisted planting operations, farmers usually steer their tractors by visually following field edges during the first row of operations. On the other hand, field edges at the headland are usually the signs that farmers use to slow down their tractors and be ready for turns. Therefore, the field edge is a useful visual reference in agricultural operations. Some parties might be interested in finding out whether it is possible to use machine vision to detect the edges of an open field, and use them to either determine the guidance direction of an agricultural vehicle, or to report the approach to a headland. This research was motivated by this intention.

Due to the wide variety of different types of open agricultural fields and edges, this research was unable to cover every possibility. As a form of preliminary exploration, some open fields with

grass-covered edges, which are often found on farms in the Midwestern United States, were used in this research. Given these conditions, the objective of this research was to investigate the possibility of using machine vision to distinguish between an open field and its grass-covered edges.

5.2 THE CONCEPT OF FIELD EDGE DETECTION USING MACHINE VISION.

5.2.1 Related Work on Field Edge Detection

Machine vision-based field edge detection has been used in remote sensing for the convenience of agricultural statistics. Several researchers (Rydberg and Borgefors, 2001; Tan et al., 2004) have used image processing techniques to divide a satellite image according to the field edges shown on it. A popular method applied by these researchers was to use a texture-based gradient edge detector to find field edges. A satellite image usually covers a localized area of fields; however, a vehicle-mounted navigation camera covers only a regional area. This difference made it difficult to adapt existing methods and apply them to this research. Therefore, new approaches had to be developed.

5.2.2 Possible Visual Clues for Field Edge Detection

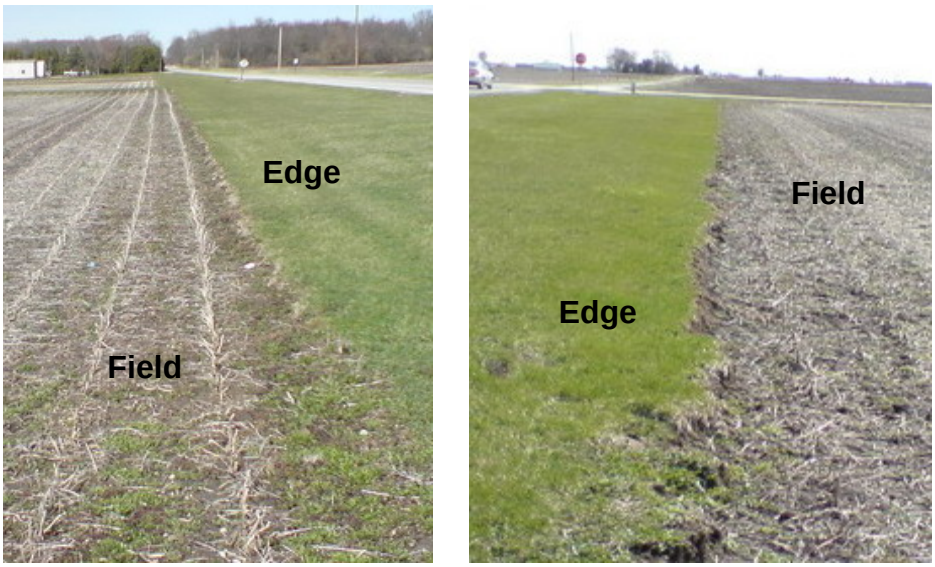
In order to detect the edge of an open field by using images obtained from an on-board camera, a number of visual clues are needed in order to form a segmentation rule for developing a machine vision algorithm. After scrutinizing the selected type of open fields (with grass-covered edges), it was found that an open field and its edge are distinguished from each other in three ways: color, texture and elevation, which are potential visual clues for field edge detection.

Color difference was the most obvious clue of the three. A top view of some open fields of the selected type is shown in Figure 54. This image shows that open fields are gray in color, but their edges are green. Furthermore, this color difference applies to all of the six pieces of fields shown in

the image, which indicates that color difference can be a global visual clue for edge detection for an entire field.

Figure 54. The top view of open fields and their edges (Urbana, IL, USA).

There were texture differences between an open field and its edge, but these differences were also found among different parts of a field. By taking a close look at an open field and its edge in Figure 55, it was found that the grass-covered edges showed a homogeneous texture (no significant patterns); while the open fields had patterns that were generated by previous farming operations and residues. This texture difference can help differentiate edges from fields. However, the patterns of open fields varied. For example, the open field in Figure 55(a) had a striped pattern due to the existence of residual crop rows, but the open field in Figure 55(b) had a non-structured pattern. Therefore, it would be difficult to determine a global segmentation rule for field edges based on texture differences.



(a) (b)

Figure 55. Visual textures of open fields and their edges.

The elevation difference between an open field and its edge was also noticed in the field scrutiny. As shown in Figure 56, the edge between a field and a ditch is significantly higher than the field. Provided that stereovision is applied, this difference can be detected by terrain analysis. The same problem, as what was experienced when texture differences were used, impeded the use of elevation difference as a global visual clue. The edges between a field and a field lane in Figure 55 do not exhibit a significant elevation difference.

The above discussion shows that color difference is the most appropriate clue to be sensed in the fields. Therefore, color was chosen to be used to find visual clues for the purpose of detecting field edges in this research.

Figure 56. Elevation difference between an open field and its edge.

5.2.3 Hue Difference Between an Open Field and Its Edge

The HSI (hue, saturation, intensity) color model was selected for use in this research for the purpose of detecting color differences. The HSI model decouples the intensity component from the color carrying information (hue and saturation) in a color image, so it is a suitable tool for developing an image processing algorithm based on color descriptions that are both natural and intuitive to humans (Gonzalez and Woods, 2002). In this model, hue is a color attribute that describes a pure color using a value between 0° and 360°, whereas saturation measures the degree to which a pure color is diluted by white light. Although the colors of open fields vary due to soil and residue conditions, they are usually far from being green, which is usually the color of the field edge. This indicates that the color difference between an open field and its edge should be a pure color difference. Therefore, hue was selected as the specific color component to be used in developing a color segmentation algorithm for the detection of field edges.

5.3 ALGORITHM DEVELOPMENT

5.3.1 Region of Interest

A camera that is set up for vehicle navigation usually looks in the heading direction of the vehicle, and its field of view may cover a large area. Figure 57 shows a sample image (320×240 pixels)

obtained from a navigation camera. It was noticed, in addition to an open field and its edge, some unrelated objects such as the sky, trees and houses in the distance were also included in the view. These unrelated objects are of no assistance in the detection of field edges, and may even introduce difficulties into the algorithm design. Therefore, a region of interest (ROI) was used to define an effective working area (the dashed line rectangle in Figure 57). Depending on the field of view that is needed for a specific application or camera installation, the range of ROI will differ; however, the definition rule does not change, which is to determine a window that excludes unrelated objects at a great distance (the top part of an image) and blurry objects that were nearby (the bottom part of an image). The objects that were nearby moved faster than other objects in the camera's view, which caused the blur. Blurry objects are of no assistance in finding a field edge, so the bottom part of the image was excluded as well.

Figure 57. Region of interest for field edge detection.

5.3.2 Hue Map

A hue map of the ROI in Figure 57 was generated in order to examine the hue attributes of an open field and its edge. The hue map is shown in Figure 58. The letters A, B, C and D correspond to the four vertices of the ROI in Figure 57. The hue map originates at the lower left corner of the ROI and represents the hue values of all of the pixels in the ROI. Some misleading peaks (those which are marked by ovals) appeared on the map. They were a kind of noise that impeded the determination of a proper segmentation rule for field edges. Therefore, a 15×15 median filter was

used to smooth the map. The smoothing result (Figure 59(a)) clearly showed that the hues of the open field and its edge were on two different levels. Figure 59(b) is the front view of the smoothed hue map, which provides a better view that allows for comparisons between the two levels. The hue values of the field pixels cluster around 30 degrees, which is in the brown color zone in the HSI model. The hue values of the edge pixels cluster around 57 degrees, which falls in the transitional zone between yellow and green in the HSI model. The difference between hue values indicates that a proper hue threshold (the dashed line in Figure 59) can be used to distinguish between the two parts. However, some significant outliers, which are marked by an oval in Figure 59(b), lay on the right hand side of the smoothed hue map. These outliers were actually from vertex C of the ROI in Figure 57, where there is a small part of a road with a color fairly close to the open field. This noise will be removed later in the algorithm.

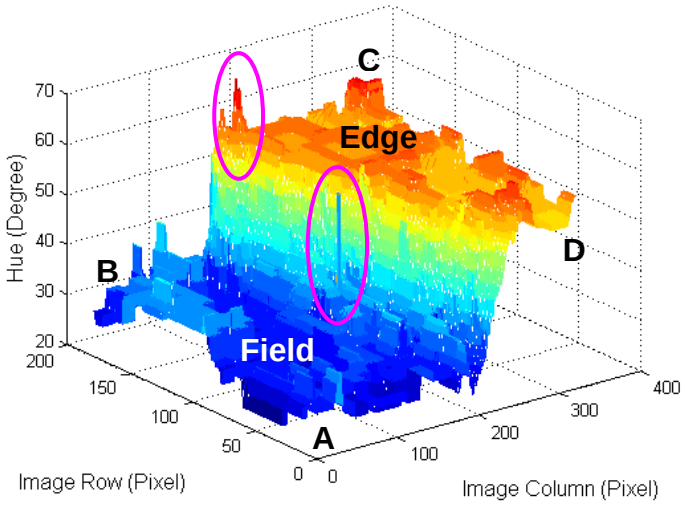
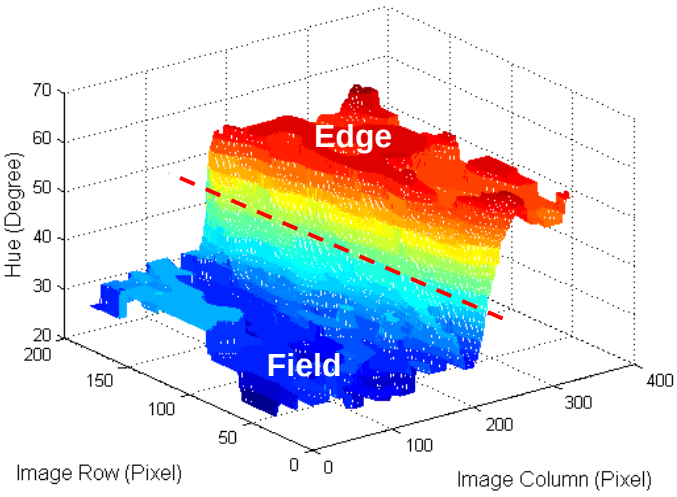
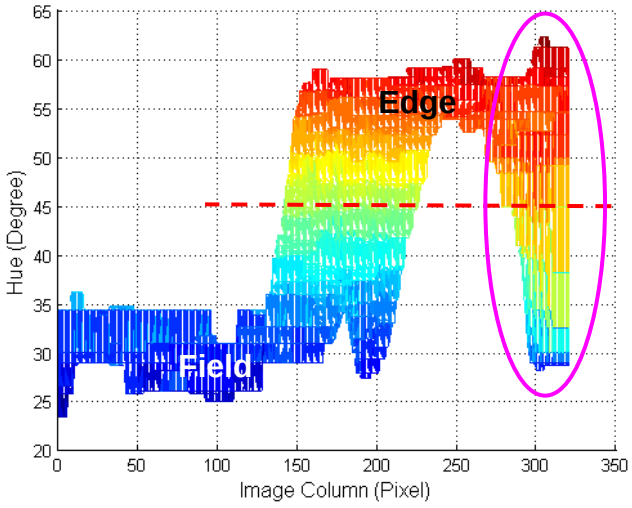


Figure 58. Hue map of the ROI.



(a)



(b)

Figure 59. Hue map after smoothing.

5.3.3 Hue Threshold

The histogram of the smoothed hue map in Figure 59 was used to determine the field-edge threshold. The histogram shows the occurrence of every hue value that was considered. Therefore, the hue value will range between 23° and 62° (an approximate reading from Figure 59(b)). Figure 60 shows the histogram, the distribution of which has two peaks: the left one corresponds with the

pixels of the field, while the right one corresponds with the pixels of the edge. Obviously, a threshold that lies in the valley between the two peaks (the dashed line in Figure 60) can be used to distinguish between these two parts. Otsu threshold selection method for histograms (Otsu, 1979) was used to determine the threshold. This method automatically determines a threshold by dividing a histogram into two major parts. As regards a histogram with two peaks, a value between the two peaks will be determined to be the threshold. In the example shown in Figure 60, a hue threshold value of 42.4° was returned by using Otsu method. After using this value to threshold the ROI, the field edge stood out in a binary image (Figure 61(b)).

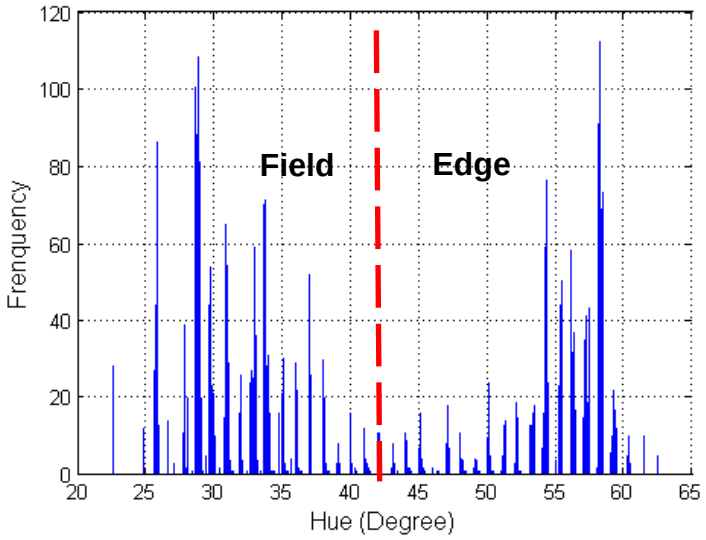
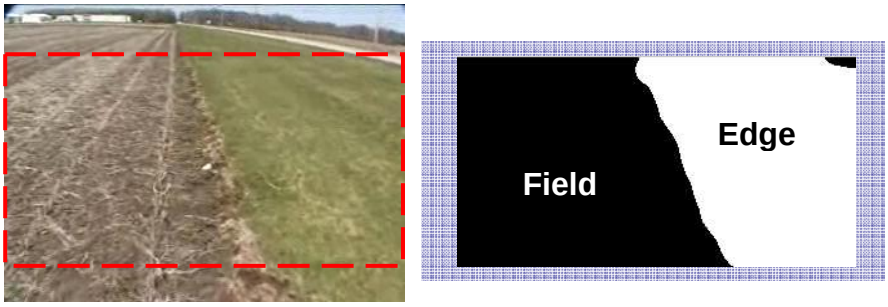


Figure 60. Histogram of the smoothed hue map.



(a) (b)

Figure 61. (a) Original image and the ROI. (b) Hue thresholding results for the ROI.

5.3.4 Field Edge Detection

An edge-detection algorithm based on gradient calculation was applied to the binary image in Figure 61(b), and afterwards the boundary between the field and its edge was enhanced (Figure 62). However, some noise appeared in the upper right corner. This visual noise was caused by the road that was included in the ROI (see vertex C of the ROI in Figure 57). In order to refine the detection result, a morphological routine was executed to remove the noise. First, the routine used a full dilation (using a 5×5 mask) to the result from the edge detection. The purpose of this operation was to bridge possible gaps along the boundary between the field and its edge. Given that the hue map was fairly smooth in the areas around the threshold value, which can be observed in the part that is close to the dashed line in Figure 59, the lengths of the possible gaps were small, usually no more than two pixels. Therefore, a 5×5 mask was large enough to bridge the gaps in a boundary. Second, the routine deleted all of the regions that had areas of fewer than 300 pixels. Noise usually had an area smaller than 300 pixels for this particular application. However, if conditions (such as image resolution) change, this thresholding value will require tuning. After applying these two steps, the noise was removed from the binary image of Figure 62 and only the clear boundary that was indicated by a bold line remained (Figure 63). In order to obtain a better visual effect, this processing result was overlapped with the original image, and the boundary between the open field and its edge was successfully highlighted (Figure 64).

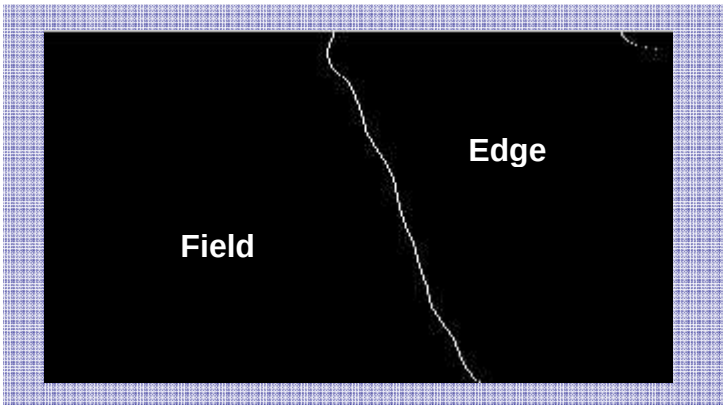


Figure 62. Edge detection results.

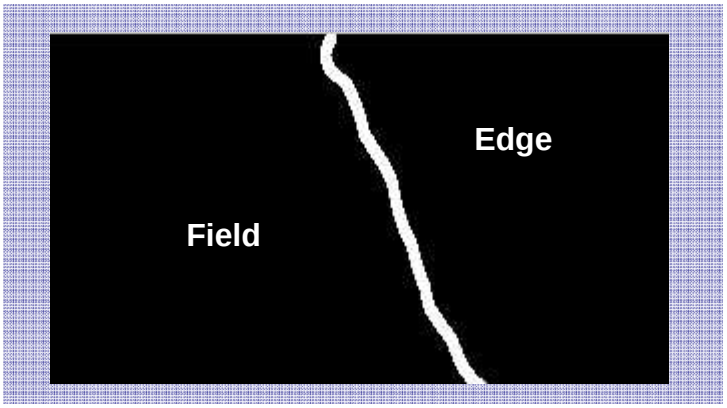


Figure 63. Field-edge boundary enhancement after noise reduction routines.

Figure 64. The detected boundary overlaid with the original image.

5.3.5 Algorithm Development Summary

The block diagram of the proposed algorithm is shown in Figure 65. Once the parameters for the ROI dimension in step 1 and the noise removing process in step 4 were set according to working conditions, the algorithm was able to process input images and highlight the boundary between an open field and its edge.

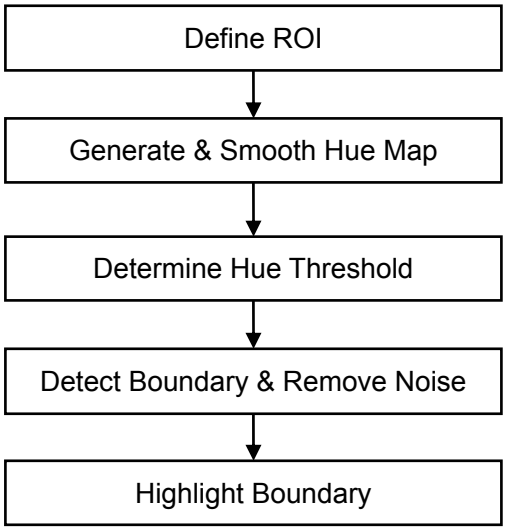


Figure 65. Block diagram for field edge detection using hue segmentation.

5.4 DESIGN OF EXPERIMENTS

Field tests were conducted at the South Farms of the University of Illinois (Urbana, IL, USA) during the spring of 2009. The experimental vehicle platform for this research was a John Deere Gator™ Utility Vehicle (Deere & Co, Moline, IL, USA) (Figure 11). A binocular stereo camera (Videre Design, Menlo Park, California, USA) was mounted on a rigid frame at the front part of the vehicle. The camera was able to acquire color images (320×240 pixels) at five frames per second, and transfer them to an onboard computer for storage and processing. Because the proposed algorithm was based on 2D image processing, only the images obtained from the left lens of the binocular camera were used. The experimental field was a rectangular open field with soybean residue, without tillage. The vehicle was driven along the four sides of the open field (Figure 66) while the vision system continuously recorded images. The camera's field of view was the same as the view shown in Figure 57. Because the research goal at the current stage was to assess the feasibility of using machine vision to detect field edges, the field test images were post processed rather than processed online. The selected ROI excluded 20% of the pixel rows from the top of each image and 20% of the pixel rows from the bottom of each image. As regards the area

thresholding value used to remove noise after the binary image dilation process, the value of 300 was kept because the image resolution was the same as the image resolution in the initially proposed algorithm.

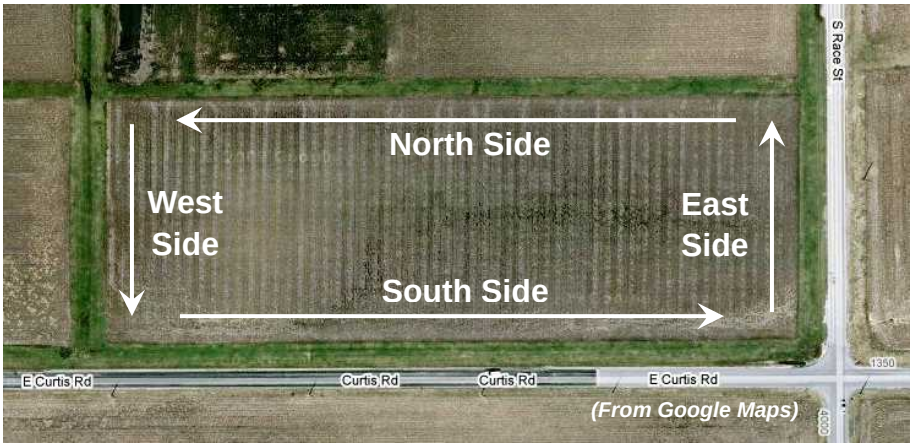


Figure 66. Experimental field used for testing the field edge detection algorithm.

5.5 RESULTS

Generally speaking, there were two types of edges in the testing field: relatively clear edges and unclear edges. Relatively clear edges had an obvious boundary between the open field area and the edge area, as shown in Figure 67(a). The east, west and south sides of the testing field (Figure 66) were relatively clear edges. As regards the unclear edge of the north side, the field and its edge merged gradually and did not have an obvious boundary, as shown in Figure 67(b). The north side and the four corners of the testing field were unclear edges.

(a) (b)

Figure 67. (a) A relatively clear edge. (b) An unclear edge.

The relatively clear edges were correctly detected by the algorithm. The detection lines followed the field edges closely (Figure 68). Although the detection lines oscillated at the lower part, the upper part maintained an acceptably consistent orientation. This difference was caused by the image resolution loss as object distance increases. An image resolution analysis was conducted by taking a picture (320×240 pixels) of two parallel straight lines with a distance of 1.5 m (Figure 69). At the bottom of the ROI (the dashed rectangle), there were 183 pixels within a distance of 1.5 m, which means that the resolution was 122 pixels per meter. However, at the top of the ROI, the resolution was only 26 pixels per meter. Therefore, the edges that were close to the camera had higher resolution and more details than did the distant edges in an image. Field edges were not perfectly straight, but with local irregularities, which were represented by higher resolution, but suppressed by lower resolution. Therefore, the lower part of the detection line (in a higher resolution part of an image) showed oscillations from image to image. On the contrary, the upper part of the detection line (in a lower resolution part of an image) showed more consistency.

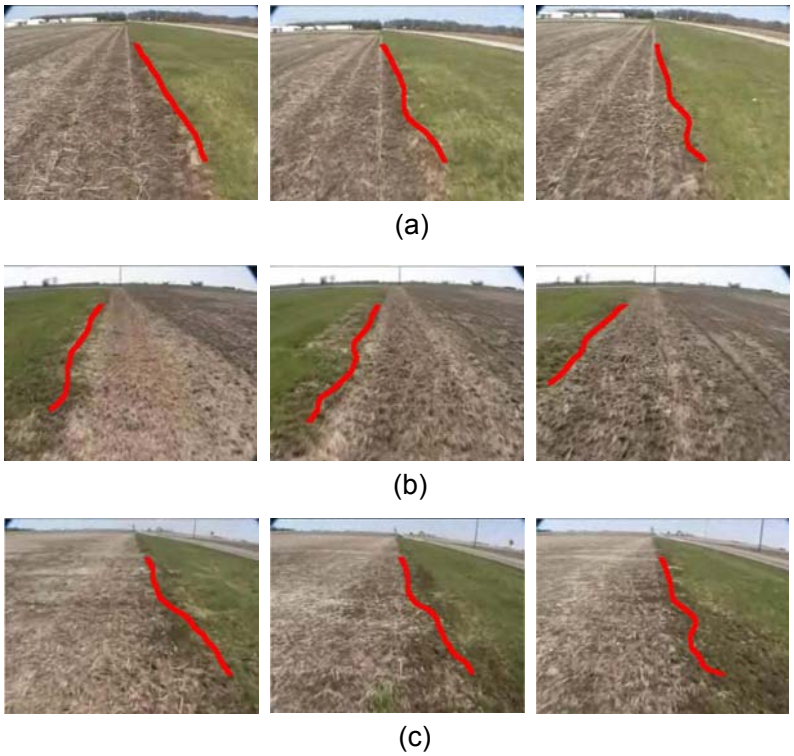


Figure 68. Detection result for relatively clear field edges. (a) East side of the testing field. (b) West side of the testing field. (c) South side of the testing field.

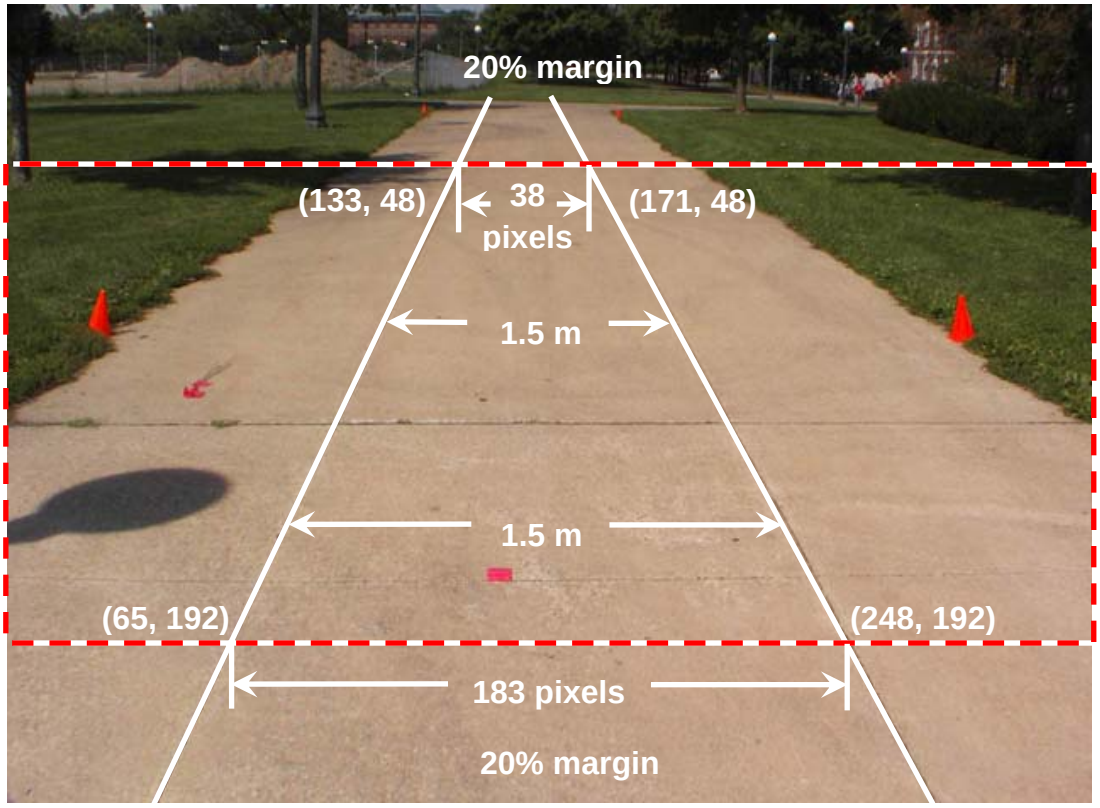


Figure 69. Image Resolution Analysis.

The developed algorithm returned inconsistent results when it was applied to the set of images with unclear edges. As shown in Figure 70, the north side (Figure 70(a)) and the vertices (Figure 70(b)) of the testing field had boundaries along the manually drawn dashed lines, but the detection results (the red lines) did not match the actual boundaries properly. The reason for this outcome was that the unclear edges exhibited the strong presence of noise, which sometimes confused the algorithm.

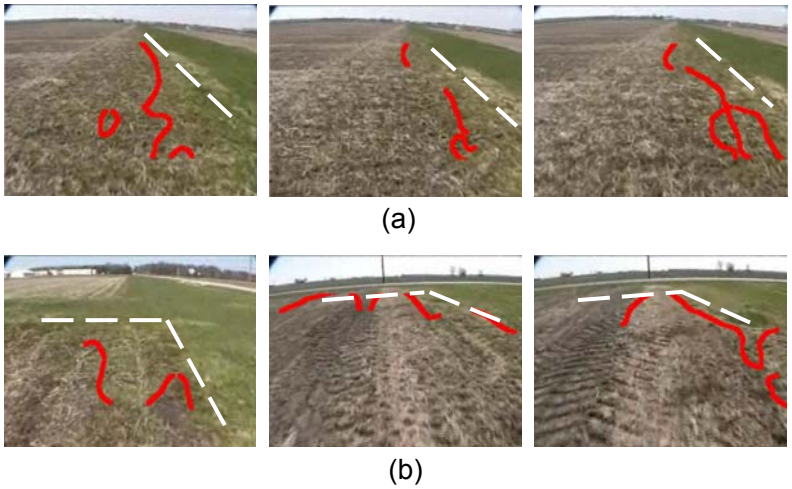


Figure 70. Detection result for unclear field edges. (a) North side of the testing field. (b) Vertices of the testing field.

5.6 DISCUSSION

According to the results of the field tests, three issues merit discussion for the purpose of improving the proposed algorithm:

(1) How to improve the detection of unclear edges.

(2) The field test dealt with a single area of field. What would happen if another field had a color that was different than the color that was found in the current testing field?

(3) How to deal with the situation where only an open field or a grass-covered edge appears in the ROI.

5.6.1 Improvements to Unclear Edges Detection

The previous section showed that the algorithm worked better when the image had lower resolution and fewer details. Therefore, image frames that had unclear edges, which were taken on the north side of the testing field, were reduced from 320×240 pixels to 160×120 pixels. On the other hand, the median filter used to remove noise on the hue map was increased from 15×15 to 30×30. Although a larger mask required more computational power, the smaller image size

compensated for the extra cost. A comparison between the two hue maps of the images with different resolutions revealed that the hue map of the low resolution image had a smoother surface than did the case for the original image (Figure 71). This proves that noise was reduced by lowering image resolution and increasing the size of the median filter. After applying the remainder of the operations of the algorithm on the smoother hue map, the detection result was greatly improved (Figure 72). Therefore, using low resolution and a large median filter helped improve the robustness of field edge detection for the proposed algorithm.

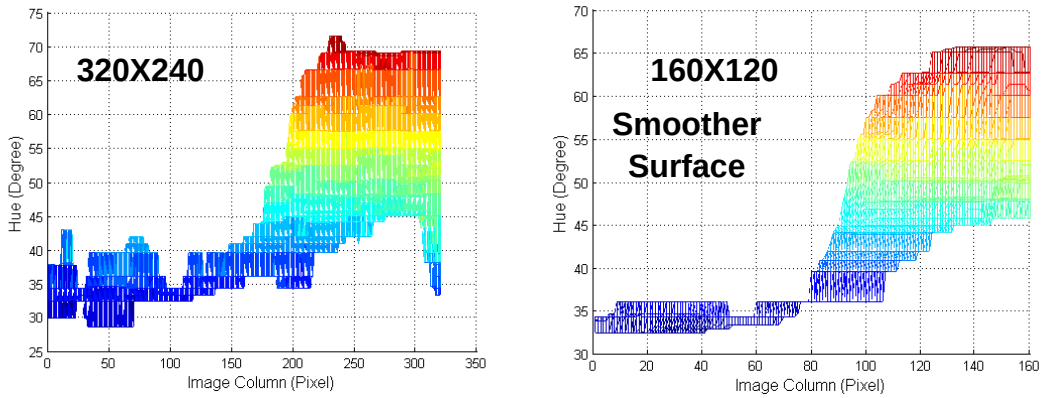


Figure 71. Comparison of the hue maps of the original image and the image with lower resolution.

Figure 72. Comparison of the detection results in the original image and the image with lower resolution.

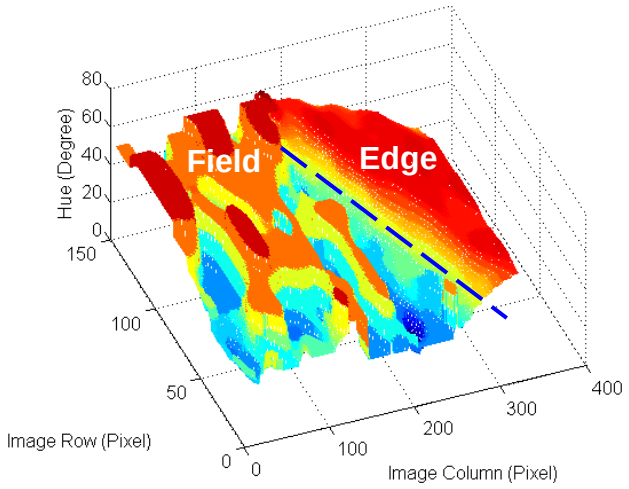
5.6.2 Color Change Adaptability

The proposed method used color as a single parameter for the purpose of distinguishing an open field and its edge. If a change in field conditions provided a significant variation in color, the algorithm might not work anymore. In order to explore the algorithm's adaptability to color changes, some images were obtained from another field, which was an open field with primary tillage and abundant corn residue (Figure 73), and these were tested in the algorithm. In comparison with the previous field, the edge looked similar, but the farming area exhibited a significantly different appearance. Due to the primary tillage, the field had a much darker color than was the case for the untilled field. There was some noticeable residue clustered in the field, which made the color of the field appear less uniform. The original algorithm was unable to deal with these changes; therefore, the detection results (red lines in Figure 73) were unacceptable.

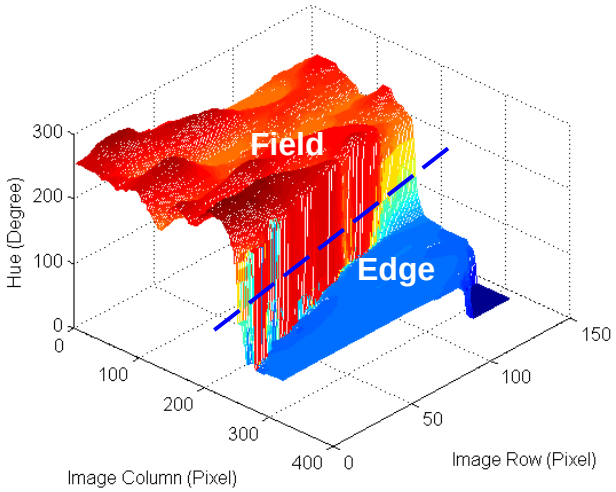
Figure 73. Field edge detection results using the initial algorithm in a new open field.

The reason for the failure was that the field and its edge could not be distinguished from each other in hue, which can be seen in the hue map of Figure 74(a). In order to help the algorithm with this situation, histogram equalization was used to enhance the hue contrast. First, the ROI (RGB image) was decomposed into its three image planes (red, green and blue); second, histogram equalization was performed on each image plane separately. In the end, the three new RGB planes were used to calculate the new image hue. Figure 74(b) shows that the hue contrast was significantly enhanced

after the histogram equalization. The hue values of the field were greater than those of the edge, the opposite of what happened in the previous cases (Figure 59). However, the original algorithm still worked for the new field and resulted in the reliable detection of field edges (Figure 75), because the hue thresholding method detected the sharp change on the hue maps.



(a)



(b)

Figure 74. Comparison of hue maps. (a) Original image. (b) Image after histogram equalization.

Figure 75. Field edge detection after histogram equalization.

Color changes can be caused by many other factors, such as natural illumination, or some other variety of farming operations that alter the soil. Therefore, having a color-adaptive capability is crucial for vision-based field edge detection, which should be a goal for future research.

5.6.3 More Intelligent Threshold Determination

The proposed algorithm determined hue thresholds dynamically. It worked well when an input image had both the open field and the grass-covered edge that showed up in the ROI, because the image hue histogram had two peaks and the Otsu threshold determination method was able to determine a hue value that was between the two peaks as a threshold. When the ROI contains only the open field or the edge, the algorithm can not be expected to report the determination of any boundary. However, the Otsu method would still return a threshold value using only one peak in the hue histogram, and this may result in an incorrect detection. This problem should be examined in the future in order to improve the robustness of the algorithm. One possible solution is to check the distribution of the hue histogram. When both the open field and the edge are in the ROI, the distribution is usually wide because the two parts have different hues. On the contrary, if only one of them is in the ROI, the distribution should be narrow. Therefore, the distribution width of the hue histogram can be used to determine whether there is a boundary for the open field and the edge that shows up in the image. If there is one, the threshold determination method will be applied to

help locate the position of the boundary.

5.7 CONCLUSIONS

A color-based algorithm was proposed for the purpose of investigating the feasibility of using machine vision to detect field edges in open field agricultural environments. This preliminary exploration used open fields with grass-covered edges, which are common in the Midwestern United States, as a study environment. The hue difference between an open field and its edge was selected for the purpose of developing the detection algorithm. The algorithm first generated a hue image of the open field and its edge, and then calculated a threshold in order to distinguish between these two parts; third, after removing noise, it highlighted the boundary between the open field and its edge. Field tests showed that the algorithm could detect relatively clear field edges, but failed to detect unclear field edges. Further upgrades of the algorithm showed that images with lower resolution helped reduce noise, and improved the detection of unclear edges. A color change adaptive function was suggested and applied in fields with different properties. The field test results demonstrated that it was possible to use machine vision to determine the boundary between an open field and its grass-covered edge. A process for checking the distribution of the hue histogram was recommended in order to improve the robustness of the algorithm.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

6.1 SUMMARY OF CONCLUSIONS

Since the mid-1980s when machine vision was initially introduced to agricultural automatic guidance, most of the research has focused on applications in structured environments. However, some farming operations (such as ploughing and seeding) occur in open fields. Open fields are an ill-structured environment, and lack visual clues, which challenges the existing methodologies and algorithms for machine vision-based agricultural vehicle navigation. In order to extend the capability of machine vision into this new application area, this dissertation involved conducting a feasibility study of machine vision applications for agricultural vehicle navigation in open field environments.

By way of a preliminary investigation, this dissertation dealt with some of the fundamental issues in vision-based agricultural vehicle navigation. These issues were: (1) camera installation pose automatic calibration; (2) vehicle heading estimation and (3) field edge detection. A stereo color camera was selected to support the research on the three issues.

An automatic calibration method was developed for a stereo camera installed on an agricultural vehicle in order to estimate its installation pose. By using a set of consecutive image frames captured by a vehicle-mounted stereo camera, the method was able to detect, and then track, natural texture features on typical agricultural grounds as references for calibrating the camera installation pose. A validation method was also developed to test the proposed calibration algorithm. Field tests showed that the calibration method could successfully estimate camera poses in the environments of relatively even soybean fields, corn fields and gravel driveways.

An image processing algorithm was proposed for the purpose of investigating the feasibility of using stereovision to estimate the heading direction of a moving vehicle in open agricultural fields. The algorithm detected static natural ground features in corner shapes, and used them as references

for estimating a vehicle heading direction. Field tests were conducted in order to evaluate the algorithm. The algorithm worked properly when the vehicle traveled in both a straight mode and an oscillating mode. However, the estimation results from straight mode showed higher accuracy than those from oscillating mode, because the algorithm design was based on the assumption that there were no camera rotations. In summary, the proposed algorithm and the field tests showed that it was possible to use stereovision to estimate the heading direction of a moving vehicle in open agricultural fields.

A color-based algorithm was proposed in order to investigate the feasibility of using machine vision to detect field edges in open field agricultural environments. As a form of preliminary exploration, open fields with grass-covered edges, which are common in the Midwestern United States, were used as a study environment. The hue difference between an open field and its edge was selected for the purpose of developing the detection algorithm. The algorithm first generated a hue image of the open field and its edge, and then it calculated a threshold in order to distinguish these two parts; third, after removing the noise, it highlighted the boundary between the open field and its edge. The field tests showed the algorithm could detect relatively clear field edges, but failed to detect unclear field edges. Further upgrades of the algorithm showed that images with lower resolution helped reduce noise, and improved the detection of unclear edges. A color change adaptive function was suggested and used in fields with different properties. The field test results demonstrated that it was possible to use machine vision to determine the boundary between an open field and its grass-covered edge.

In conclusion, this research proved the feasibility of the machine vision applications with respect to the three primary issues. Therefore, machine vision has the navigational capability to properly direct the motions of agricultural vehicles in open field environments.

6.2 RECOMMENDATIONS FOR FUTURE RESEARCH

(1) Camera rotations

The algorithms for both the yaw estimation in camera installation pose calibration and the heading direction estimation were based on the assumption that the vehicle (camera) engaged only in translational movements. In real applications, camera rotations are inevitable due to field bumpiness or vehicle turns. In order to improve the accuracy of camera pose calibration and vehicle heading estimation, the effect of camera rotations is recommended to be included in the two estimations.

(2) High accuracy positioning sensors for heading estimation validation

The absence of high accuracy positioning sensors prevented the realization of a more accurate evaluation of the heading estimation. As regards future research, it is recommended that high accuracy GPS and IMU be used to record vehicle movements during field tests, so that benchmarks for heading estimation validation can be obtained.

(3) Real-time processing

This research applied post processing to all three of the applications for the purpose of preliminary evaluation. As regards practical uses, camera installation pose calibration may not require real-time processing because calibrations are not conducted frequently, but vehicle heading estimation and field edge detection do require real-time processing. Therefore, efforts should be made to realize real-time processing for these two applications in the future.

(4) Hue histogram distribution checking for more robust field edge detection

As mentioned in section 5.6.3, when the ROI contains only the open field or the edge, the algorithm would incorrectly report detections of the boundary between a field and its edge. One possible solution to this problem is to check the distribution of the hue histogram, which is a recommendation for future research. When both the open field and the edge are in the ROI, the distribution is usually wide because the two parts have different hues; on the contrary, if only one

of them is in the ROI, the distribution should be narrow. Therefore, the distribution width of the hue histogram can be used to determine whether there is a boundary between the open field and the edge that shows in the image. If there is a boundary, the threshold determination method should be applied in order to help locate the boundary position.

(5) Machine vision-based positioning for agricultural vehicles

This research proved that natural ground features in an open agricultural field can be detected and tracked in 3D space. Therefore, the relative motion of a ground vehicle with respect to these static features can be obtained. By using the idea of dead reckoning, connecting the relative motions from each step reconstructs the vehicle trajectory, which helps localize the vehicle in an open field. This machine vision-based positioning idea for agricultural vehicles represents a possible direction for future research.

REFERENCES

Agrawal, M., K. Konolige, and R. C. Bolles. 2007. Localization and mapping for autonomous navigation in outdoor terrains: A stereo vision approach. In *Proc. Workshop on Applications of Computer Vision*, 6-13. Austin, TX, USA: IEEE.

Angelova, A., L. Matthies, D. Helmick, and P. Perona. 2007. Learning and prediction of slip from visual information. *Journal of Field Robotics* 24(3): 205-31.

Barawid Jr., O. C., A. Mizushima, K. Ishii, and N. Noguchi. 2007. Development of an autonomous navigation system using a two-dimensional laser scanner in an orchard application. *Biosystems Engineering* 96(2): 139-49.

Bellutta, P., R. Manduchi, L. Matthies, K. Owens, and A. Rankin. 2000. Terrain perception for DEMO III. In *Proc. IV 2000 Intelligent Vehicles Symposium*, 326-31. Dearborn, MI, USA: IEEE.

Benson, E. R., J. F. Reid, and Q. Zhang. 2000(a). Development of an automated combine guidance system. ASAE Paper No. 003137. St. Joseph, MI: ASAE.

Benson, E. R., J. F. Reid, Q. Zhang, and F. A. C. Pinto. 2000(b). An adaptive fuzzy crop edge detection method for machine vision. ASAE Paper No. 001019. St. Joseph, MI: ASAE.

Biesiadecki, J. J., P. Chris Leger, and M. W. Maimone. 2007. Tradeoffs between directed and autonomous driving on the mars exploration rovers. *International Journal of Robotics Research* 26(1): 91-104.

Billingsley, J., and M. Schoenfisch. 1995. Vision-guidance of agricultural vehicles. *Autonomous Robots* 2(1): 65-76.

Cheng, Y., M. W. Maimone, and L. Matthies.. 2005. Visual odometry on the Mars exploration rovers. In *Proc. The International Conf. on System, Man and Cybernetics*, 903-10. Waikoloa, HI, USA: IEEE.

Cheng, Y., M. W. Maimone, and L. Matthies. 2006. Visual odometry on the Mars exploration rovers: A tool to ensure accurate driving and science imaging. *IEEE Robotics and Automation Magazine* 13(2): 54-62.

Collado, J. M., C. Hilario, A. de la Escalera, and J. M. Armingol. 2006. Self-calibration of an on-board stereo-vision system for driver assistance systems. In *Proc. Intelligent vehicles symposium*, 156-62. IEEE.

Constable, G., and B. Somerville. 2003. *A century of innovation: Twenty engineering achievements that transformed our lives*. Washington D.C., USA: Joseph Henry Press.

Coulombeau, P., and C. Laurgeau. 2002. Vehicle yaw, pitch, roll and 3D lane shape recovery by vision. *Intelligent Vehicle Symposium* 2(2): 619-25.

Fischler, M. A., and R. C. Bolles. 1980. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* 24(6): 381-395.

Franke, U., D. Gavrila, S. Gorzig, F. Lindner, F. Paetzold, and Ch Wohler. 1998. Autonomous driving goes downtown. *IEEE Intelligent Systems & their Applications* 13(6): 40-8.

Gennery, D. B. 1980. Modelling the environment of an exploring vehicle by means of stereo vision. Ph.D. dissertation. Stanford, CA: Stanford University, Department of Computer Science.

Gonzalez, R. C., and R. E. Woods. 2002. *Digital image processing*. Upper Saddle River, NJ, USA: Prentice Hall.

Han, S., Q. Zhang, B. Ni, and J. F. Reid. 2004. A guidance directrix approach to vision-based vehicle guidance systems. *Computers and Electronics in Agriculture* 43(3): 179-95.

Harris, C., and M. Stephens. 1988. A combined corner and edge detection. In *Proc. The Fourth Alvey Vision Conf.*, 147-51. Manchester, UK: BMVA.

Keicher, R., and H. Seufert. 2000. Automatic guidance for agricultural vehicles in Europe. *Computers and Electronics in Agriculture* 25(1): 169-94.

Kise, M., and N. Noguchi. 2005. Unmanned robot tractor uses RTK-GPS and attitude sensor. *Resource: Engineering and Technology for Sustainable World* 12(7): 7-8.

Kise, M., Q. Zhang, and F. Rovira Mas. 2005. A stereovision-based crop row detection method for tractor-automated guidance. *Biosystems Engineering* 90(4): 357-67.

Konolige, K. 1998. Small vision systems: Hardware and implementation. In *Proc. Eighth International Symposium on Robotics Research*, 203-12. Shonan, Japan: IFRR.

Konolige, K., M. Agrawal, R. C. Bolles, C. Cowan, M. Fischler, and B. Gerkey. 2008. Outdoor mapping and navigation using stereo vision. *Springer Tracts in Advanced Robotics* 39: 179-90.

Krotkov, E., S. Fish, L. Jacket, B. McBride, M. Perschbacher, and J. Pippine. 2007. The DARPA PerceptOR evaluation experiments. *Autonomous Robots* 22(1): 19-35.

Labayrade, R., D. Aubert, and J. P Tarel. 2003. Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation. In *Proc. Intelligent Vehicle Symposium*, 646-51. Piscataway, NJ, USA: IEEE.

Lamprecht, B., S. Rass, S. Fuchs, and K. Kyamakya. 2007. Extrinsic camera calibration for an on-board two-camera system without overlapping field of view. In *Proc. 10th International IEEE Conf. on Intelligent Transportation Systems*, 265-70. Piscataway, NJ, USA: IEEE.

Leemans, V., and M. -F Destain. 2006. Application of the Hough transform for seed row localisation using machine vision. *Biosystems Engineering* 94(3): 325-36.

Li, J., and N. M. Allinson. 2008. A comprehensive review of current local features for computer vision. *Neurocomputing* 71(6): 1771-87.

Ma, Y., S. Soatto, J. Kosecka, and S. Sastry. 2005. *An invitation to 3-D vision*. New York, NY, USA: Springer.

Marchant, J. A. 1996. Tracking of row structure in three crops using image analysis. *Computers and Electronics in Agriculture* 15(2): 161-79.

Marita, T., F. Oniga, S. Nedevschi, T. Graf, and R. Schmidt. 2006. Camera calibration method for far range stereovision sensors used in vehicles. In *Proc. IEEE Intelligent Vehicles Symposium*, 356-63. Piscataway, NJ, USA: IEEE..

Matthies, L. 1992. Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation. *International Journal of Computer Vision* 8(1): 71-91.

Matthies, L., and S. M. Shafer. 1987. Error modeling in stereo navigation. *Robotics and Automatio*n 3(3): 239-48.

Matthies, L., M. Maimone, A. Johnson, Y. Cheng, R. Willson, C. Villalpando, S. Goldberg, A. Huertas, A. Stein, and A. Angelova. 2007. Computer vision on mars. *International Journal of Computer Vision* 75(1): 67-92.

Moravec, H. P. 1983. The Stanford cart and the CMU rover. *Proceedings of the IEEE* 71(7): 872-84.

Nister, D., O. Naroditsky, and J. Bergen. 2006. Visual odometry for ground vehicle applications. *Journal of Field Robotics* 23(1): 3-20.

O'Connor, M., G. Elkaim, and B. Parkinson. 1995. Kinematic GPS for closed-loop control of farm and construction vehicles. In *Proc. the 8th International Technical Meeting of the Satellite Division of the Institute of Navigation. Part 1 (of 2)*, 1261-68. Palm Springs, CA, USA: Inst. of Navigation.

Ollis, M., and A. Stentz. 1997. Vision-based perception for an automated harvester. In *Proc. IROS '97*, 1838-44. Grenoble, France: IEEE.

Otsu, N. 1979. A threshold selection method from gray-level histograms. *Systems, Man and Cybernetics, IEEE Transactions* 9(1): 62-6.

Pinto, F. A. C., and J. Reid. 1998. Heading angle and offset determination using principal component analysis. ASAE Paper No. 983113. St. Joseph, MI: ASAE.

Pomerleau, D., and T. Jochem. 1996. Rapidly adapting machine vision for automated vehicle steering. *IEEE Expert* 11(2): 19-27.

Reid, J., and S. Searcy. 1987. Vision-based guidance of an agriculture tractor. *Control Systems Magazine* 7(2): 39-43.

Reid, J. F., Q. Zhang, N. Noguchi, and M. Dickson. 2000. Agricultural automatic guidance research in North America. *Computers and Electronics in Agriculture* 25(1): 155-67.

Rovira-Más, F., Q. Wang, and Q. Zhang. 2009. Bifocal Stereoscopic Vision for Intelligent Vehicles. *International Journal of Vehicular Technology*, Article ID 123231, 9 p.

Rovira-Más, F., Q. Zhang, J. F. Reid, and J. D. Will. 2005. Hough-transform-based vision algorithm for crop row detection of an automated agricultural vehicle. *Journal of Automobile Engineering* 219(8): 999-1010.

Rydberg, A., and G. Borgefors. 2001. Integrated method for boundary delineation of agricultural fields in multispectral satellite images. *IEEE Transactions on Geoscience and Remote Sensing* 39(11): 2514-20.

Sappa, A. D., D. Geronimo, F. Dornaika, and A. Lopez. 2006. On-board camera extrinsic parameter estimation. *Electronics Letters* 42(13): 745-7.

Subramanian, V., T. F. Burks, and A. A. Arroyo. 2006. Development of machine vision and laser radar based autonomous vehicle guidance systems for citrus grove navigation. *Computers and Electronics in Agriculture* 53(2): 130-43.

Tan, Q., J. Hu, S. Bi, and Z. Liu. 2004. A study on rice field edge extraction in radarsat SAR images. In *Proc. 2004 IEEE International Geoscience and Remote Sensing Conf.*, 4239-42. Piscataway, NJ, USA: IEEE.

Thorpe, C., M. Herbert, T. Kanade, and S. Shafer. 1991. Toward autonomous driving: the CMU navlab. Part I: perception. *IEEE Expert* 6(4): 31-42.

Ting, K. C., and T. Grift. 2005. Automation for bioprocessing and bioproduction. *Resource: Engineering and Technology for Sustainable World* 12(8): 29.

Weber, J. W., and M. Atkin. 1997. Further results on the use of binocular vision for highway driving. In *Proc. Transportation Sensors and Controls: Collision Avoidance, Traffic Management, and ITS*, 52-61. Boston, MA, USA: SPIE.